

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DE 8 MAI 1945 - GUELMA -
FACULTÉ DES MATHÉMATIQUES, D'INFORMATIQUE ET DES SCIENCES DE LA MATIÈRE

Département d'Informatique



Mémoire de Fin d'études Master

Filière : Informatique

Option : Système informatique

Thème _____

**FindToME : Un Système de Recommandation
Collaboratif de Lieux**

Encadré Par :

DR. BOUGHAREB Djalila

Présenté par :

GUERRI Nesrine

Juin 2022

Remerciements

Je voudrais dans un premier temps remercier, mon encadreur de mémoire DR.Boughareb, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je remercie mes très chers parents pour leur soutien inestimable et surtout mon père pour sa précieuse aide à la relecture et à la correction de mon mémoire.

Un grand merci à mon mari pour ses conseils et son soutien inconditionnel.

Je remercie ma sœur Malak et mon frère Amine, pour leur encouragements.

je remercie également mes amis imen et hazem qui ont toujours été là pour moi.

DÉDICACE

À MES CHERS PARENTS

Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma Considération pour les sacrifices que vous avez consenti pour mon instruction et Mon bien être. Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours.

Que ce modeste travail soit l'exaucement de vos vœux tant formulés, le fruit de vos Innombrables sacrifices, bien que je ne vous en acquitterai jamais assez.

Puisse Dieu, le Très Haut, vous accorder santé, bonheur et longue vie et faire en sorte que Jamais je ne vous déçoive.

RÉSUMÉ

Ce projet se positionne dans le cadre du domaine des systèmes de recommandation (SR).

Un système de recommandation est un outil de recherche d'information et de filtrage qui vise à proposer aux utilisateurs des items pouvant l'intéresser. Globalement, les SR se basent sur l'analyse des préférences des utilisateurs et leurs évaluations implicites ou explicites pour les items.

Rechercher et trouver un bon médecin, restaurant, école, dentiste, cafeteria, etc. en temps réel à partir de sa position actuelle reste difficile à réaliser en Algérie. Dans le but de faciliter à l'utilisateur algérien l'accès aux différents lieux. Nous avons développé un système de recommandation collaboratif appelé FindToMe qui apprend à classer les différents lieux selon leurs distances et les avis des utilisateurs.

Une étape de collecte d'avis et de feedback utilisateurs a été menée sur les différents lieux au niveau de la ville de Guelma. Puis, nous avons appliqué une analyse des données collectées en se basant sur la méthode d'analyse de sentiments VADER. Ensuite, pour la recommandation des lieux intéressants un algorithme de recommandation est conçu et implémenté en fonction des quatre (04) paramètres suivants : la localisation, les notes d'évaluations (rating), les commentaires et l'âge.

Les résultats que nous avons obtenus sont encourageants et nous permettent de conclure que notre système de recommandation réussit à suggérer des items qui correspondent

aux préférences des utilisateurs.

Mots Clés : systèmes de recommandation, lieux, géo localisation, filtrage collaboratif, analyse des sentiments.

ABSTRACT

This project is positioned in the field of recommender systems (RS). A recommendation system is an information search and filtering tool that aims to propose to users items that propose to users items that may be of interest. Globally, RS are based on the analysis of users preferences and their implicit or explicit evaluations for items.

Searching for and finding a good doctor, restaurant, school, dentist, cafeteria, etc. in real time from one's current from one's current location remains difficult to achieve in Algeria. In order to facilitate to the Algerian user access to different places. We have developed a collaborative recommendation system called FindToMe which learns to classify the different places according to their distances and the opinions of users reviews. A step of collecting users opinions and feedback has been conducted on the different places in the city of Guelma. Then, we applied an analysis of the data collected based on the VADER sentiment analysis method.

Then, for the recommendation of the interesting places a recommendation algorithm is designed and implemented according to the following four (04) parameters : the location, the rating, the comments and the age. The results we obtained are encouraging and allow us to conclude that that our recommendation system succeeds in suggesting items that correspond to the users preferences.

Keywords : recommender systems, places, geo-location, collaborative filtering, sentiment analysis.

TABLE DES MATIÈRES

Liste des figures	xii
Liste des tableaux	xiii
1 Chapitre 01 : les systèmes de recommandation	1
1.1 Introduction	1
1.2 Définition	1
1.3 Types de système de recommandation :	3
1.3.1 Systèmes de recommandation basés sur le contenu(CBRS) :	5
1.3.2 Systèmes de recommandation collaboratifs :	5
a- Méthode basée sur la mémoire (MBCF)	6
Filtrage basé sur les utilisateurs :	7
- Calcul de la similarité :	7
Filtrage basé sur les items :	9
b- Méthode basée sur le modèle (MCF)	9
c- La différence entre les deux méthodes :	10
1.3.3 Systèmes de recommandation hybrides (HRS) :	11
1.4 Autres classifications :	13
1.4.1 Systèmes de recommandation basés sur la connaissance (KBRS) :	13

1.4.2	Systèmes de recommandation basés sur les ontologies (OBRS) :	14
1.4.3	Systèmes de recommandation intelligents (IRS) :	14
1.4.4	Systèmes de recommandation basés sur la démographie(DBRS) :	15
1.4.5	Systèmes de recommandation sensibles au contexte (CARS) : .	15
	a- Systèmes de recommandation sensibles au temps (TiARS) : .	15
	b- Systèmes de recommandation basés sur les avis (RBRS) : . .	16
1.5	La recommandation et l'analyse des sentiments :	16
1.5.1	Approches pour mesurer le sentiment :	16
1.5.2	L'approche lexical :	17
1.5.3	L'approche d'apprentissage automatique :	17
	L'approche hybride :	18
1.6	Méthode d'analyse des sentiments « VADER Valence AwareDictionary and sentiment Reasoner :	19
1.7	Prédiction de lieux et systèmes de recommandation :	20
1.8	Conclusion	21
2	Chapitre 02 : Conception	22
2.1	Introduction	22
2.2	Méthodologie	22
2.2.1	Collection des données :	22
2.2.2	Prétraitement :	23
2.2.3	Analyse des sentiments :	23
	Score positif :	24
	Score négatif :	25
2.2.4	Importation de la Base de données :	25
2.3	Fonctionnement du système de recommandation	26
2.4	Conception du site FindToMe :	28
2.4.1	Diagramme de cas d'utilisation « FindToMe »	28
2.4.2	Diagramme de classe « FindToMe » :	29

2.4.3	Diagramme de séquence :	30
	Diagramme de séquence « Login »	31
	Diagramme de séquence « Inscription »	32
	Diagramme de séquence « ajouter au favori »	33
2.5	Algorithme de recommandation	33
	Algorithme FindToMe :	33
2.6	Conclusion	35
3	Chapitre 03 : Implémentation	36
3.1	Introduction	36
3.2	Environnements de développement et langages utilisés	36
3.2.1	PHP (Hypertext Preprocessor) :	37
3.2.2	Sublime text editor :	37
3.2.3	HTML (Hypertext Markup Language), CSS (Cascade StyleSheet) :	37
3.2.4	Leaflet :	37
3.2.5	Apify :	38
3.2.6	JavaScript :	38
3.2.7	MySQL :	39
3.2.8	Bootstrap :	39
3.2.9	AJAX :	39
3.2.10	GOOGLECOLAB :	39
3.2.11	Python :	40
3.2.12	NumPy :	40
3.2.13	Pandas :	40
3.2.14	VADER (Valence Aware Dictionary for Sentiment Reasoning) :	41
3.2.15	JQuery :	41
3.2.16	AppsGeyser :	41
3.3	Les interfaces graphiques du système (Findtome) :	42
3.3.1	Espace Administrateur :	42

3.3.2	Espace utilisateur :	45
	Exemple de recommandation d'une cafétéria :	50
3.4	Résultat et discussion :	53
3.5	Conclusion	56

TABLE DES FIGURES

1.1 catégorisation des systèmes de recommandation d'après Lashkmi et Bavhani 2021 [29]	4
1.2 Filtrage collaboratif basé sur la mémoire	7
1.3 Techniques d'hybridation [8]	12
2.1 Un exemple d'une partie de la collection des données qui concerne les restaurants.	23
2.2 Un exemple réel positif depuis la base.	24
2.3 Un exemple réel négatif depuis la base.	25
2.4 conception de la base de données.	26
2.5 Fonctionnement du système FindToMe	27
2.6 Diagramme de cas d'utilisation « FindToMe »	29
2.7 Diagramme de classe « FindToMe »	30
2.8 Diagramme de séquence « Login »	31
2.9 Diagramme de séquence « Inscription »	32
2.10 Diagramme de séquence « ajouter aux favoris »	33
3.1 Espace administrateur	42
3.2 la liste des utilisateurs du système "findtome".	43
3.3 l'ensemble de données présentées par le dashboard.	44

3.4	Profil administrateur.	45
3.5	Connexion au système.	45
3.6	Formulaire d'inscription du système.	47
3.7	Page d'accueil du système	48
3.8	l'ensemble des catégories disponibles sur le système.	48
3.9	le champ dédié aux commentaires des différents utilisateurs.	49
3.10	la rubrique dédiée aux notations par étoiles.	49
3.11	Espace d'historique des utilisateurs.	50
3.12	Recommandation du système.	50
3.13	Le choix de lieu.	51
3.14	Détails de lieu.	51
3.15	Commentaire positif.	52
3.16	le résultat obtenu après le commentaire.	52
3.17	Recommandation automatique.	53

LISTE DES TABLEAUX

3.1	La moyenne des évaluations et des précisions des utilisateurs de système "Findtome".	54
3.2	Calcul de précision pour l'utilisateur numéro 26.	55

INTRODUCTION GÉNÉRALE

Ces dernières années, le développement technologique et notamment l'internet et les réseaux sociaux ont incité les chercheurs à explorer les systèmes de recommandation (RS) pour générer des services personnalisés. L'adoption généralisée du web en tant que plateforme commerciale a entraîné des changements fondamentaux dans les habitudes des personnes (clients- commerçants) ce qui a généré une grande concurrence et une surcharge d'information.

Pour pallier à ces difficultés, des systèmes de recommandation ont vu le jour entant qu'outil d'aide à la prise de décision en temps réel dans divers domaines tel que : le commerce électronique, les voyages, le tourisme, les films et l'apprentissage en ligne, etc.

Les systèmes de recommandation apparaissent comme une approche populaire et fiable de filtrage d'informations, capable de suggérer des articles, et des lieux pertinents à l'utilisateur cible et actif en fonction de ses préférences et ses intérêts dynamiques. De différents algorithmes de recommandation ont été appliqués pour fournir un moyen automatique et intelligent permettant de filtrer l'excès d'informations disponibles et de faire des recommandations personnalisées de produits et de services lors d'une interaction directe.

Parmi les techniques développées pour faire des recommandations personnalisées nous pouvons citer : le filtrage collaboratif, le filtrage basé sur le contenu et le filtrage

hybride. Ces techniques ont été largement utilisées dans le but de fournir à l'utilisateur une liste personnalisée de recommandations.

Rechercher et trouver un bon médecin, restaurant, école, dentiste, cafeteria, etc. en temps réel à partir de sa position actuelle reste difficile à réaliser en Algérie. En effet, peu d'utilisateurs expriment leurs avis sur un lieu ou un service sous forme de commentaires, de j'aime/je n'aime pas ou d'étoiles. En conséquence, le manque de feedback rend impossible de déterminer la qualité d'un service donné et influence ainsi la concurrence et le progrès.

Dans le but de faciliter à l'utilisateur algérien l'accès aux différents lieux et de bénéficier d'un service de plus en plus qualifié, ce projet consiste à développer un système de recommandation collaboratif appelé FindToMe qui apprend à recommander les différents lieux selon leurs distances et les avis des utilisateurs, l'avis de l'utilisateur courant s'il a déjà visité et évalué le lieu et l'âge. Une étape de collecte d'avis et de feedback utilisateurs a été menée sur les différents lieux (restaurant, etc.) au niveau de la wilaya de Guelma. Puis, nous avons opté pour appliquer une analyse des données collectées en se basant sur la méthode d'analyse des sentiments VADER. Finalement, FindToMe affiche des détails sur leurs adresses et les évaluations données par les visiteurs de ces lieux.

Le système reçoit les feedbacks des utilisateurs sur les recommandations données dans le but d'améliorer d'avantage les recommandations ultérieures.

Ce mémoire est composé de 3 chapitres :

Le premier chapitre définit le concept du système de recommandation, et présente les différentes approches de recommandation ainsi qu'une synthèse des travaux effectués dans cet axe.

Le deuxième chapitre est consacré aux différentes phases de conception de notre système à savoir : la collecte de données, le prétraitement en utilisant une méthode d'analyse de sentiments, l'apprentissage et le test en appliquant l'algorithme de recommandation proposé et finalement, la métrique utilisée pour la validation des résultats de recommandation.

Le troisième chapitre présente les détails d'implémentation du système et présente également les résultats obtenus et les évaluations élaborées.

CHAPITRE 1

CHAPITRE 01 : LES SYSTÈMES DE RECOMMANDATION

1.1 Introduction

Dans ce chapitre nous allons présenter des généralités sur les systèmes de recommandation, leurs types ainsi que leurs domaines d'application. Nous abordons également de façon générale les méthodes d'analyse de sentiments en mettant l'accent sur la méthode VADER que nous avons appliqué dans ce projet. Une synthèse des travaux existants est aussi présentée à la fin de ce chapitre.

1.2 Définition

Plusieurs définitions ont été donné pour les systèmes de recommandation. (Jie et al., 2015) [52] les considèrent comme des outils qui aident les utilisateurs à la surcharge d'informations et à l'amélioration de la gestion de la relation client tout en fournissant des recommandations (Produits ou services personnalisés) aux utilisateurs.

D'autres les définissent comme une sous-classe des systèmes de filtrage d'informations. Ces systèmes prédisent les "notes" ou "préférences" que les utilisateurs attribuent à un produit en recueillant des informations sur les préférences des utilisateurs pour un ensemble d'éléments[42].

(Adomavicius et al., 2005) [4] considèrent que les systèmes de recommandation prédisent les évaluations pour les produits inconnus, à chaque utilisateur, par l'utilisation des évaluations des autres utilisateurs, et recommande les N meilleurs items ayant la plus haute valeur d'évaluation prédite.

Robin Burke [12] qui précise qu'ils s'agissent de « systèmes capables de fournir des recommandations personnalisées permettant de guider l'utilisateur vers des ressources intéressantes et utiles au sein d'un espace de données important ».

Récemment Anwar et Uma (2019c, 2020a) [8] disent qu'un système de recommandation a un rôle important pour surmonter le problème de la surcharge d'informations et trouver des produits et des services qui correspondent à nos préférences personnelles.

De ce qui précède, nous pouvons conclure que Les systèmes de recommandation sont des outils de filtrage des informations dont l'objectif essentiel est de fournir à l'utilisateur des recommandations personnalisées facilitant le traitement de grandes quantités de données dans un temps record.

Les deux entités essentielles dans un système de recommandation sont l'item et l'utilisateur.

Item : c'est l'entité qui représente les besoins recherchés (liste de recommandation).

Utilisateur : c'est la personne pour laquelle le système recommande des items en se basant sur ses préférences et ses traces. Le domaine d'information d'un système de recommandation contient aussi une évaluation exprimée par un utilisateur pour un item, appelée note, vote, préférence ou une évaluation. Exprimée en valeur numérique dans une échelle quelconque, la plus utilisée est (l'intervalle [1-5]) ou binaire (aimer Ne pas aimer, bon , etc.). Ces expressions sont stockées dans une matrice,

qu'on note R (pour Rating) où : $R : U \times I$. On note l'évaluation donnée par un utilisateur u à un item i par $r_{u,i}$ (le rating donné par l'utilisateur u à l'item i) qui interprète le triplet (u,i,r) où une note de 5, représente une grande préférence et une note de 1 représente une faible préférence (l'utilisateur n'a pas aimé l'item).

1.3 Types de système de recommandation :

Pour classer les SRs en plusieurs catégories, plusieurs méthodes de filtrage ont été proposées, Récemment, Lakshmi et Bhavani [29] proposent de classer les techniques de recommandation en deux grandes classes :

- La recommandation orientée vers une perspective du domaine : elle dépend de la recommandation de produits, de vidéos des articles ou citations et la recommandation dans la santé.
- La recommandation orientée vers une perspective algorithmique qui explore plusieurs volets théoriques comme : le filtrage basé contenu, collaboratif, hybride, filtrage basé sur les graphes et le filtrage sensible au contexte. Détaillé comme suit : (voir figure (01) ci-dessous)

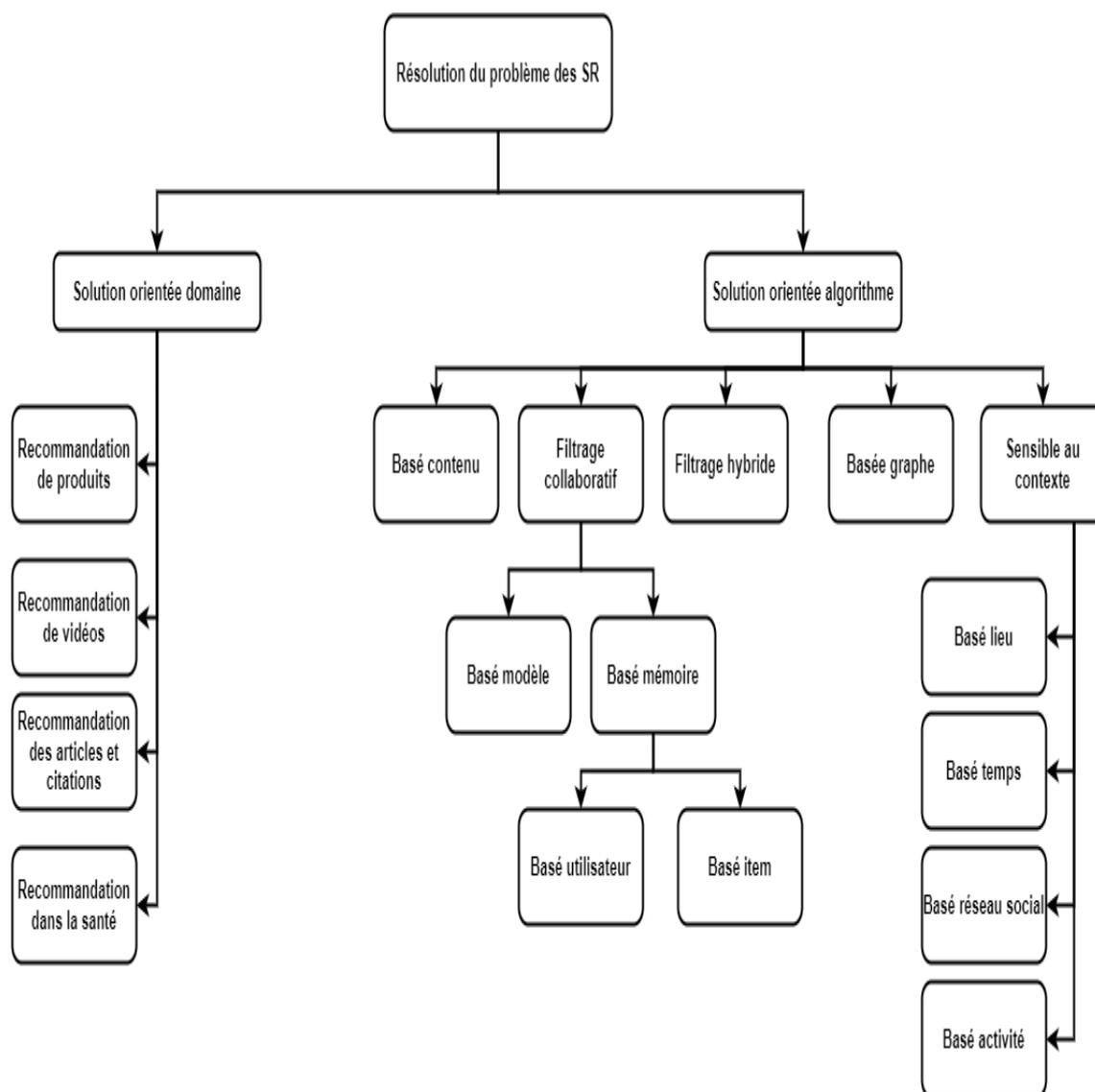


FIGURE 1.1 – catégorisation des systèmes de recommandation d'après Lashkmi et Bavhani 2021 [29]

En effet, la topologie traditionnelle établie par (Adomavicius et al.,2005) qui considèrent trois types de systèmes de recommandation : systèmes basés sur le filtrage Collaboratif (FC), le filtrage basé sur le Contenu (FBC) et le filtrage hybride qui consiste en une hybridation des forces et des faiblesses du FC et FBC. Dans ce qui suit, nous allons détailler ces trois classes.

1.3.1 Systèmes de recommandation basés sur le contenu(CBRS) :

CBRS fonctionne essentiellement sur l'analyse des caractéristiques d'un article préféré par l'utilisateur et détermine les similitudes entre les différents articles.

La Préférence des utilisateurs est calculée en fonction de ses choix passés. Les Eléments similaires sont identifiés par rapport aux caractéristiques des éléments comparés [37]. Le système peut recommander des articles à condition que les informations sur ces articles sont disponibles. Pour une recommandation appropriée dans le CBRS, beaucoup d'informations sont requises pour les éléments.

Un CBRS lui-même peut être divisé en approches basées sur les attributs [18] et en approches de raisonnement basé sur le cas [4]. Les méthodes basées sur les attributs sont suggérées dans la base de la correspondance entre leurs caractéristiques et les profils des apprenants, et les méthodes de raisonnement par cas proposent des éléments liés aux intérêts des utilisateurs. Généralement, les méthodes CBRS utilisent principalement des modèles de récupération simples comme la pondération TermFrequency-Inverse Document Frequency (TF-IDF), la correspondance des mots clés et le modèle d'espace vectoriel (VSM) [8].

1.3.2 Systèmes de recommandation collaboratifs :

Le filtrage collaboratif (CF) est la méthode la plus pratiquée dans le monde de la recommandation. Elle est basée sur l'hypothèse "stéréotypée", c'est-à-dire un utilisateur faisant partie d'un groupe de personnes ayant eu des préférences similaires dans le passé [15].

Un CF essaie de prédire les évaluations basées sur les éléments d'un utilisateur du même type à propos des préférences passées des autres utilisateurs qui sont similaires. Cette méthode permet de faire des recommandations sans connaître les caractéristiques des items ou des utilisateurs, ce qui rend son utilisation plus simple que celle basé sur le filtrage basé sur le contenu. (Breese, 1998) [19] fournit un classement accepté aujourd'hui, qui divise les méthodes de recommandation basées sur le filtrage

collaboratif en deux catégories :

- 1) À base de mémoire (appelée aussi à base de : heuristique, voisinage proche)
- 2) À base de modèle.

a- Méthode basée sur la mémoire (MBCF)

Les méthodes basées sur la mémoire sont celles qui ne fonctionnent que sur la matrice d'évaluation des utilisateurs pour les items (utilisateur x items) et utilisent n'importe quelle évaluation (rating/préférence) générée avant le traitement des recommandations à condition que la matrice doit être mise à jour régulièrement avant la compilation. L'utilisateur est un vecteur dont les caractéristiques sont des items et dont les évaluations (implicites ou explicites) sont des valeurs données aux caractéristiques.

Evaluations explicites : Ce sont des données non interprétées, fournit directement par l'utilisateur.

Evaluations implicites : ce sont des données qui ne demandent pas d'intervention par l'utilisateur, elles sont recueillies suite à la surveillance des actions des utilisateurs sur la plateforme qu'ils utilisent [15].

La figure suivante résume la collaboration à base de mémoire.

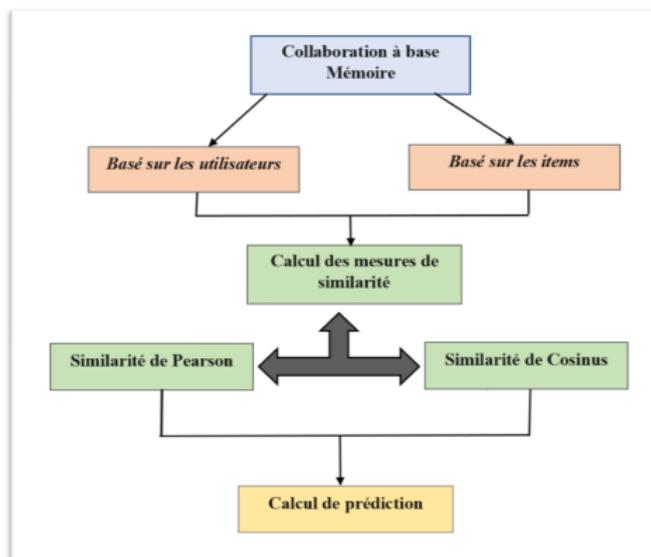


FIGURE 1.2 – Filtrage collaboratif basé sur la mémoire

On explique ci-dessous le calcul de la similarité et de la prédiction pour les deux types de filtrage (basé sur les utilisateurs et basé sur les items).

Filtrage basé sur les utilisateurs :

Le filtrage collaboratif basé sur les utilisateurs a été introduit pour la première fois dans le système GroupLens [41] , Il s'agit de déterminer les utilisateurs qui sont similaires à l'utilisateur courant, ensuite calculer une valeur de prédiction pour chaque item candidat à la recommandation en analysant les évaluations que les voisins de l'utilisateur courant ont exprimées sur cet item.

- Calcul de la similarité :

La similarité entre deux utilisateurs u et v est calculée en utilisant la similarité Cosinus ou en utilisant le coefficient de corrélation de Pearson Selon Schafer et al 2007 [13], à noter que le coefficient de Pearson est le plus utilisé compte tenu de sa performance en terme de pertinence des recommandations.

La similarité de cosinus :

Cosinus est une mesure de similarité entre deux objets a et b de manière globale, très pratiquée en recherche d'informations [43] , qui consiste à représenter les deux objets par deux vecteurs $V_u, V_{u'}$ et de mesurer le cosinus de l'angle formé par les deux vecteurs.

$$\text{sim}(u, u') = \cos(v_u, v_{u'}) = \frac{\sum_{i \in I_{uu'}} r_{u,i} \times r_{u',i}}{\sqrt{\sum_{i \in I_{uu'}} r_{u,i}^2} \cdot \sqrt{\sum_{i \in I_{uu'}} r_{u',i}^2}}$$

[1]. Où v_u et $r_{u'}$ représentent les vecteurs des utilisateurs u et u' respectivement et il s'agit de : $r_{u,i}$ et $r_{u',i}$ respectivement $I_{uu'}$ consiste en l'ensemble des items notés à la fois par l'utilisateur u et l'utilisateur u' .

la corrélation de pearson :

Ce coefficient est spécialement utilisé pour calculer la similarité entre deux utilisateurs u et v par les auteurs du système GroupLens [41]. Le coefficient de corrélation de Pearson mesure la relation entre la covariance et le produit de l'écart type des scores donnés par deux utilisateurs, de ce fait il utilise les items notés à la fois par u et v pour mesurer la similarité.

$$\text{sim}(u, u') = \text{Pearson}(u, u') = \frac{\sum_{i \in I_{uu'}} (r_{u,i} - \bar{r}_u) \cdot (r_{u',i} - \bar{r}_{u'})}{\sqrt{\sum_{i \in I_{uu'}} (r_{u,i} - \bar{r}_u)^2} \cdot \sqrt{\sum_{i \in I_{uu'}} (r_{u',i} - \bar{r}_{u'})^2}}$$

[2].

Où \bar{r}_u et $\bar{r}_{u'}$ sont les moyennes des notes données par les utilisateurs u et u' respectivement sur les items qu'ils ont notés.

calcul de la prédiction :

La prédiction p de note qu'un utilisateur u peut donner à un item i est donnée par la formule suivante :

$$p = \bar{r}_u + \frac{\sum_{k \in N(u) \cap u_i} \text{sim}(k, u) \times (r_{k,i} - \bar{r}_k)}{\sum_{k \in N(u) \cap u_i} |\text{sim}(k, u)|}$$

[3].

Où $N(u)$ est l'ensemble des voisins de l'utilisateur u et U_i est l'ensemble des utilisateurs ayant notés l'item i .

Filtrage basé sur les items :

Le filtrage basé sur les items est le contraire du filtrage basé sur l'utilisateur, on calcule la similarité entre les items, comme la similarité de cosinus entre deux items i et j est représentée par la formule suivante (équation 4) :

$$\text{sim}(i, j) = \cos(v_i, v_j) = \frac{\sum_{u \in U_{ij}} r_{u,i} \times r_{u,j}}{\sqrt{\sum_{u \in U_{ij}} r_{u,i}^2} \cdot \sqrt{\sum_{u \in U_{ij}} r_{u,j}^2}}$$

[4].

Où U_{ij} est l'ensemble des utilisateurs ayant notés les items i et j .

calcul de la prédiction :

La formule de prédiction de la note qu'un utilisateur u peut donner à un item j est comme suit (équation 5) :

$$(u, j) = \frac{\sum_{i \in I_{u(i)}} \text{sim}(i, j) \times r_{u,i}}{\sum_{j \in I_{u(i)}} |\text{sim}(i, j)|}$$

[5].

b- Méthode basée sur le modèle (MCF)

Parmi les modèles les plus pratiqués, il y'a les classificateurs bayésiens, les réseaux de neurones, les systèmes de logiques floues (fuzzy systems), les algorithmes génétiques, les factorisations de matrice (matrix factorization), les facteurs latents (latent features).

Ces techniques donnent des recommandations en évaluant les différents paramètres

statistiques.

Premièrement, les évaluations passées des utilisateurs pour des items sont rassemblées et utilisées pour définir un modèle. Ce dernier est utilisé pour faire des prévisions d'appréciation qui sont plus rapides et plus précises [4].

Les algorithmes de cette méthode peuvent être divisés en trois sous-types : clustering, factorisation matricielle (MF) et l'apprentissage en profondeur (Deep Learning).

Clustering (KNN) : Le principe de cette approche est le même que celle des systèmes de recommandation basés sur la mémoire. Dans les algorithmes basés sur la mémoire, on utilise les similarités entre les utilisateurs et / ou les items comme pondérations pour prédire une note pour un utilisateur et un item. La différence est que les similitudes dans cette approche sont calculées sur la base d'un modèle d'apprentissage non supervisé, plutôt que sur la corrélation de Pearson ou la similarité cosinus.

Dans cette méthode, il est impérieux de limiter le nombre d'utilisateurs similaires à k , ce qui rend le système plus évolutif[3].

Factorisation matricielle (MF) : les préférences d'un utilisateur peuvent être déterminées par un petit nombre de facteurs cachés appelés incorporations (Embedding) [3].

Deep Learning : Les caractéristiques latentes de l'utilisateur et les caractéristiques latentes de l'item sont recherchées à partir des matrices de notes pour une combinaison spécifique d'item-utilisateur. Elles présentent les valeurs d'entrée pour les autres couches linéaires et non linéaires. Cette entrée est passée à plusieurs couches, linéaires ou sigmoïdes et le réseau va apprendre les poids correspondants par n'importe quel algorithme d'optimisation[53].

c- La différence entre les deux méthodes :

Basée sur la mémoire :

- Coûte chère (utilisation de toutes les données).
- Utilisation importante de l'espace de stockage.

Basée sur le modèle :

- Utilise les paramètres définis grâce à l'apprentissage.
- Les prédictions se font grâce au modèle préalablement appris d'où l'utilisation bien moins de données, en conséquence l'obtention des résultats plus rapide [15].

1.3.3 Systèmes de recommandation hybrides (HRS) :

Il est nécessaire de combiner différentes méthodes pour avoir des systèmes plus performants en recommandation.

La méthode HRS combine les caractéristiques de plusieurs méthodes de filtrage comme les méthodes CBRS et CRS en y ajoutant les avantages de chacune d'elles pour obtenir de meilleures performances [21].

Il atténue également diverses difficultés des RSs comme :

La montée en puissance : c'est la progression spectaculaire.

La surspécialisation : le système recommande que des items similaires au profil de l'utilisateur. L'utilisateur ne peut donc recevoir que des recommandations proches des articles qu'il a notés ou remarqués dans le passé [4].

La rareté des données : signifie que les données sont peu nombreuses par rapport à la quantité requise.

Nous pouvons conclure que la méthode HRS est très utile et acceptée parce qu'elle réduit les limites des systèmes de recommandation conventionnels [8]. En 2002, (Burke, Robin. 2002) [12] a proposé un classement de sept techniques d'hybridation. Il a divisé la technique d'hybridation en sept types.

Ceci est illustré à la Figure suivante :

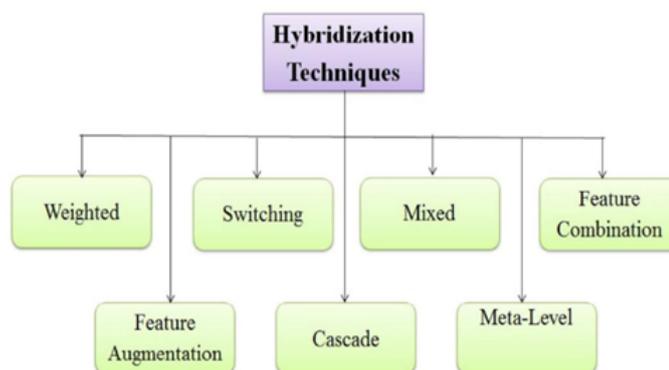


FIGURE 1.3 – Techniques d'hybridation [8]

- **Weighted** : Cette technique utilise plusieurs approches de recommandation individuellement. La recommandation des items est distribuée de façon équitable à chaque élément existant dans chaque liste de recommandation produite, et choisit l'élément avec le score le plus élevé pour la recommandation finale [15].
- **Switching** : cette technique choisit un système dynamique selon un critère lié au type de recommandation que nous voulons produire (par exemple : la recommandation à court terme vs recommandation à long terme). Cette norme peut être difficile à déterminer, donc les méthodes sont également sélectionnées en fonction de critères de confiance de la recommandation [15].
- **Mixed** : cette technique utilise plusieurs approches de recommandation en parallèle et présente un assemblage des items recommandés par les différentes méthodes. La sélection des items candidats se fait en demandant à chaque système de délivrer à ses candidats une note associée, une note prédite et/ou un indice de confiance. Ensuite, un module spécifique effectue un mélange de ces recommandations avec un tri et une sélection basée sur les scores associés aux éléments candidats. Cette méthode d'hybridation est seulement citée par (Burke, 2002) et n'est pas évaluée. [15].
- **Feature combination** : cette approche utilise des données généralement pratiquées pour un type de système de recommandation dans un autre contexte. Comme l'utilisation des données des avis des utilisateurs sur les éléments généralement traités par un système basé sur le contenu [15].

- **Feature augmentation** : cette approche rajoute des données caractéristiques des utilisateurs et des items avant de les utiliser comme données d'entrée pour le système de recommandation [15].
- **Cascade** : est une méthode hiérarchique, où chaque méthode de SR raffine une recommandation obtenue par une méthode de recommandation utilisée avant. Par exemple, le système EntreC (Burke, 2002), qui fournit trop d'items avec des scores identiques, a été mis à jour en ajoutant la post-classification basée sur la recommandation collaborative [15].
- **Meta Level** : La méthode prend en entrée le modèle fait par un autre SR, par rapport au feature augmentation, cette approche nécessite le remplacement complet du modèle d'entrée par la sortie de la recommandation précédente [15].

1.4 Autres classifications :

Les typologies citées ci-dessus ont été la base de plusieurs autres classifications de sous-systèmes de recommandation comme les systèmes de recommandation basés sur la connaissance (KBRS), les systèmes de recommandation basés sur les ontologies (OBRS), systèmes de recommandation intelligents (IRS), Systèmes de recommandation basés sur la démographie (DBRS), systèmes de recommandation sensibles au contexte (CARS), systèmes de recommandation sensibles au temps (TiARS), et les systèmes de recommandation basés sur les avis (RBRS).

1.4.1 Systèmes de recommandation basés sur la connaissance (KBRS) :

Ces systèmes fonctionnent sur la base de la connaissance du domaine de l'utilisateur et les éléments concernant l'intérêt de l'utilisateur [46]. On y trouve trois types d'informations, à savoir : la connaissance des éléments, des utilisateurs et de la similarité entre les utilisateurs et les éléments. Les approches KBRS sont utilisées pour des domaines complexes où les articles ne sont pas achetés régulièrement.

Parmi les avantages de ce type des systèmes, nous pouvons citer l'absence de problèmes de démarrage à froid. La principale limitation de KBRS est la nécessité d'avoir des connaissances et de compétences en ingénierie [42].

1.4.2 Systèmes de recommandation basés sur les ontologies (OBRS) :

Les OBRS fonctionnent sur la base d'une compréhension idéale de l'utilisateur, et des connaissances des éléments et du domaine. Dans un OBRS, l'ontologie est essentiellement utilisée pour la représentation des connaissances [20].

Les OBRS sont principalement liés aux KBRS. Ces dernières utilisent une ontologie pour la représentation des connaissances. L'OBRS rend beaucoup plus facile et éclairci certaines limitations liées aux SR traditionnels, telles que la rareté des données [33] la ramp up et la surspécialisation. Dans le même temps, l'OBRS s'appuie davantage sur l'évaluation du rapport de domaine.

Cette particularité rend les OBRS plus utiles et plus adaptés aux domaines du e-commerce, E-learning et du tourisme. D'autre part, développer des ontologies dans ce sens est un processus complexe et coûteux [8].

1.4.3 Systèmes de recommandation intelligents (IRS) :

Les IRSs sont une extension de la RS basée sur les connaissances. Elle explore les connaissances, l'apprentissage, l'exploration de nouvelles actualités d'après les avis et elle identifie des préférences. L'IRS est décrit par les composants suivants :

- 1) Méthodes d'apprentissage.
- 2) Modèles de représentation des connaissances.
- 3) Mécanisme de raisonnement.

De plus, l'IRS a cinq modèles de connaissances dans un contexte spécifique, tels que les éléments, les utilisateurs, le contexte, les commentaires et les domaines pouvant être pris en compte au moment de la recommandation.

Un IRS utilise les informations sur les utilisateurs via un mécanisme d'apprentissage pour augmenter ses performances [8].

1.4.4 Systèmes de recommandation basés sur la démographie(DBRS) :

Le principe de ces systèmes est l'identification des utilisateurs qui sont démographiquement similaires à l'utilisateur cible afin de prédire les évaluations des items. Plus précisément, un DBRS fonctionne sur la base des données démographiques des utilisateurs.

L'objectif principal de la recommandation s'agit de classer les utilisateurs sur la base des attributs et des données démographiques des utilisateurs stockés dans leur profil (le sexe, l'âge, le lieu, etc.).

Contrairement au CRS et au CBRS, la méthode démographique n'a pas besoin d'avis précédents des utilisateurs [8].

1.4.5 Systèmes de recommandation sensibles au contexte (CARS) :

Les CARSs utilisent des conditions contextuelles pour faire des meilleures recommandations aux utilisateurs [5].

Les Préférences de l'utilisateur peuvent varier selon l'environnement (humeur, saison, heure, jour, lieu, compagnon, occasion, etc.). Par exemple :

regarder un film (lieu,heure,compagnon,etc.), voyager (lieu, heure, météo, conditions de circulation, etc.) et écouter de la musique (heure, lieu, humeur, occasion, etc.) [47].

Ces systèmes peuvent être classés eux-mêmes en :

a- Systèmes de recommandation sensibles au temps (TiARS) :

Ce derniers se focalisent principalement sur l'emploi du temps comme information contextuelle. Le profit des utilisateurs peut varier ou évoluer de temps à autre et

cela est principalement pris en compte dans ce type de système de recommandation [8].

b- Systèmes de recommandation basés sur les avis (RBRS) :

C'est des systèmes de recommandation ayant l'aptitude de diminuer ainsi que d'éliminer les problèmes liés au démarrage à froid et au manque des données. Un RBRS agit sur le fondement d'un profil de produit ainsi que d'un profil d'utilisateur qui sont analysés.

Ceci, permet d'examiner et de trier les avis à l'aide de différentes méthodes telles que la recherche et l'inspection d'opinions et l'analyse de texte. La recommandation d'item est effectuée d'une manière identique à celle des CBRs et CRSs [8].

1.5 La recommandation et l'analyse des sentiments :

L'analyse de sentiments - aussi appelée opinion mining ou sentiment analysis - est un champ d'études de développement du traitement automatique du langage (TAL). Elle consiste à identifier le sentiment, l'opinion ou l'évaluation - positif ou négatif - exprimé-e-s à l'intérieur d'une unité informationnelle (i.e. mot, phrase, paragraphe ou document) [30].

Nous pouvons dire que l'analyse des sentiments est le processus qui consiste à déterminer par machine si une écriture est positive, négative ou neutre dérivant de l'opinion d'un orateur.

1.5.1 Approches pour mesurer le sentiment :

Il y a pratiquement deux approches pour qualifier le sentiment dans les textes : l'approche lexicale et l'approche statistique ou par apprentissage automatique. Une

troisième, hybride consiste en une combinaison de composants linguistiques et de modules de classification.

1.5.2 L'approche lexical :

Basée sur les dictionnaires, elle repose sur des lexiques. Un dictionnaire prédéfini dont chaque terme est associé à une polarité : principalement positif, négatif et dans certains cas neutres. Il s'agit pour la plupart de verbes et d'adjectifs mais aussi de noms communs) et d'adverbes.

Cette méthode est avantageuse dans la mesure où elle assure la transparence des critères de classification. Il est simple et efficace de rajouter, modifier ou de supprimer des termes aux lexiques pour améliorer les résultats. Toutefois, la construction manuelle de tels lexiques nécessite un travail approfondi avec une couverture relativement limitée. Les termes doivent nécessairement être listés pour être repérés. D'un autre côté, les lexiques d'analyse des sentiments sont souvent limités à des termes uniques (appelé uniterme ou uni gramme) [30].

1.5.3 L'approche d'apprentissage automatique :

L'apprentissage automatique (en anglais Machine Learning) est un sous-domaine de l'intelligence artificielle, qui donne à un système une capacité de compréhension grâce à ses algorithmes. L'idée est d'apprendre des algorithmes à partir des données et de faire des prédictions avec ces données et par cela les ordinateurs apprennent à résoudre des tâches spécifiques, sans avoir besoin de les programmer [2].

Il existe 3 types ou catégories de l'apprentissage automatique [2] :

1) Apprentissage supervisé : après avoir présenté les données et les résultats souhaités aux ordinateurs, ils auront la capacité de faire des prédictions pour de nouvelles données d'entrée.

2) Apprentissage non supervisé : on donne à l'ordinateur que les données et il doit

trouver une structure avec un sens par lui-même sans l'intervention d'une supervision extérieure. Elle dépend également du clustering. Le processus d'apprentissage se fait par l'observation et non par l'exemple.

3) Apprentissage par renforcement : c'est l'apprentissage où la machine se comporte comme un agent qui apprend de son environnement d'une manière interactive jusqu'à ce qu'il découvre les comportements qui produisent des récompenses. .

Nous pouvons dire que l'avantage de cette approche est que les valeurs de la polarité du lexique sont essentiellement calculées par estimation statistique. Cela réduit le travail humain pour la création des ressources et il est avantageux en terme de couverture.

À noter que si les textes qui sont donnés en apprentissage ne sont pas représentatifs du texte auquel le modèle prédictif est appliqué, les prévisions peuvent s'avérer inexactes [39].

L'approche hybride :

L'approche hybride comprend l'apprentissage automatique et la méthode basée sur le lexique (éventuellement au sein des règles linguistiques). Différents classifieurs de sentiments fondés sur un lexique ou sur les méthodes d'apprentissage sont utilisés en cascade, si un classifieur échoue, le suivant prend un tour pour classer, et ainsi de suite jusqu'à ce que le texte restant soit classifié. Une approche hybride peut utiliser des techniques d'apprentissage automatique pour proposer des nouveaux termes et construire des dictionnaires qui seront plus riches et plus spécifiques pour un terme dans son contexte. Les modèles hybrides donnent souvent de bons résultats en terme de précision, citons : (Mudinas et al., n.d.) avec 89.64%, (Bahrainian et Dengel, n.d.) avec 89.13%, (Zhang et al., 2015) avec 85.40%, (Malandrakis et al., 2013) avec 85.80% (Shoukry et Rafea, n.d.) avec 80.90%. Par exemple (Gupta et Joshi, 2019) présente une analyse de sentiments sur Twitter grâce à une approche hybride dans laquelle le vecteur de caractéristiques basé sur SentiWordNet (SWN) sert d'entrée au modèle de

classification SVM. Leurs résultats expérimentaux montrent que le vecteur de caractéristiques basé sur le contexte obtenu conduit à une amélioration de 2 à 6% du système d'analyse des sentiments sur Twitter [30].

1.6 Méthode d'analyse des sentiments « VADER Valence Aware Dictionary and sentiment Reasoner :

Le Vader (Valence Aware Dictionary and sEntiment Reasoner) est un outil d'analyse des sentiments basé sur des règles, spécifiquement adapté aux sentiments exprimés dans les médias sociaux.

Il est entièrement open-source sous la Licence MIT. Il est développé par Hutto et ses collègues en 2014, il prend dans une chaîne et renvoie un dictionnaire de scores dans chacune des quatre catégories : Négatif, Neutre, Positif, composé (en normalisant les scores) [24].

Cet outil a été utilisé dans de nombreux systèmes et présente de nombreux avantages. En effet, les émotions sont reconnues comme des indicateurs pour exprimer les besoins.

Tout de même, lorsqu'une personne ressent une émotion positive, cela signifie que son besoin est satisfait. Par contre lorsque la personne ressent une émotion négative, cela signifie que son besoin n'est pas satisfait [49].

Vader donne un score composé indiquant si une phrase dans le sens positif, négatif ou neutre par l'addition des scores de valence de chaque mot dans le lexique, ajuste selon les règles puis se normalise entre -1 et +1.

Cette méthode calcule un score composé en utilisant des ratios à partir des portions de texte dans chaque catégorie.

- Positif > 0,05
- Neutre entre -0,05 et 0,05
- Négatif <= -0,05 [26].

Et cette méthode a été choisie pour être utilisée dans notre application.

1.7 Prédiction de lieux et systèmes de recommandation :

Plusieurs systèmes de recommandation ont été développés dans le but de fournir des suggestions personnalisées aux utilisateurs. Ces systèmes visent à extraire (explicitement ou implicitement) les préférences des utilisateurs et à présenter des recommandations plus conformes à leurs préférences.

La disponibilité des technologies de détection de localisation telles que le GPS, a permis aux réseaux sociaux de fonctionner prenant en compte les données de localisation de différentes manières. En effet, les utilisateurs peuvent partager leurs positions sur les sites Web, télécharger des photos géo-localisées sur les plateformes de médias sociaux ou rédiger un commentaire sur un endroit visité.

Compte tenu de l'importance, des commentaires des utilisateurs sur les lieux qu'ils ont visités, de nombreux systèmes de recommandation se sont appuyé sur l'analyse des commentaires [55].

De plus, la prise en compte du contexte joue un rôle majeur dans l'amélioration de la qualité des systèmes de recommandation touristiques [1].

Parmi de nombreuses informations contextuelles, la localisation représente l'élément le plus important utilisé dans les systèmes de recommandation touristiques actuels. (Abowd et al., 1999 ; Bao, Fang et Zhang, 2014 ; Xiaolin Zheng, Ding, Lin, et Chen, 2016 ; Yochum, Chang, Gu, amp ; Zhu, 2020). À cet égard, les comportements des personnes qui se rendent dans les zones protégées du parc marin de Ningaloo, en Australie, ont été extraits à l'aide d'un SIG [44] .

De plus, la position actuelle de l'utilisateur a été utilisée pour fournir des recommandations [48]. Ces recommandations sont affichées en fonction des préférences de l'utilisateur pour les transports en commun et à pied. En plus, P*SiS* [7] suggère à l'utilisateur des destinations appropriées en fonction de son contexte, notamment : le lieu, l'heure, la vitesse, la direction et la météo. Dans ce système, l'historique touristique des utilisateurs et leurs préférences sont utilisés pour fournir des recommandations plus précises. La majeure partie des systèmes de recommandation touristiques

traditionnels se base sur les évaluations des utilisateurs concernant les lieux visités. Comme le texte peut contenir plus d'informations qu'une note, l'analyse des commentaires des utilisateurs peut améliorer la précision de ces systèmes [55]. Les chercheurs ont analysé les opinions des utilisateurs par des techniques d'exploration de texte [54].

Aussi, Loh et al 2003 ont conçu une page de discussion privée qui pose des questions aux utilisateurs pour trouver du vocabulaire spécifique dans le texte. L'objectif est de détecter un vocabulaire spécifique à une ontologie. Puis le type souhaité par l'utilisateur est extrait sur la base des concepts de l'ontologie et des recommandations sont fournies en conséquence [32]. POST-VIA360 [14] est un système de recommandation bio-inspiré développé dans le but de fidéliser les touristes après leur première visite d'une attraction. Pour se faire les auteurs utilisent une ontologie du tourisme pour fournir des suggestions basées sur les visites précédentes, l'emplacement actuel et les aspects sociaux.

Recemment, Vijesh (2021) a proposé une approche de recommandation qui exploite les technologies de l'internet des objets (internet of things) IoT pour [51].

1.8 Conclusion

Dans ce chapitre nous avons étudié les systèmes de recommandation ainsi que leurs types, leurs domaines d'application, ensuite on a parlé de l'analyse des sentiments et la méthode d'analyse des sentiments Vader que nous avons utilisé sur notre application et on a terminé ce chapitre par quelques travaux sur la prédiction de lieux appliquée dans les systèmes de recommandation.

CHAPITRE 2

CHAPITRE 02 : CONCEPTION

2.1 Introduction

L'objectif de ce chapitre est de présenter les étapes de conception de notre système de recommandation. Ainsi, nous allons présenter une description détaillée de chacune des phases suivantes : la collecte de données, l'analyse des sentiments et le principe de la recommandation. Le fonctionnement général du système est également présenté.

2.2 Méthodologie

Dans cette section, nous présentons les différentes étapes de conception de notre système.

2.2.1 Collection des données :

Nous avons collecté notre ensemble de données de plusieurs lieux (restaurants, médecins, avocat, cyberspace, etc.) à partir de Google Maps, et nous avons sélectionné comme wilaya notre ville Guelma afin d'améliorer la recommandation des

lieux comme le montre la figure suivante, où chaque lieu est décrit par le nom et la catégorie, ainsi que la position (latitude et longitude).

	name	categoryName	lat	long	positiveScores	negativesScores	rating
0	Fast food le Parisien	Restaurant	36.464339	7.432207	0.00	0.0	4.0
1	La Tulipe	Restaurant	36.454229	7.434595	18.30	0.0	4.6
2	le prince	Restaurant	36.453937	7.433425	50.00	0.0	5.0
3	Info Plus Cyber Cafe	Internet cafe	36.465415	7.428113	0.00	0.0	NaN
4	GIAPPONE Sushi Shop	Restaurant	36.462048	7.432705	0.00	0.0	5.0
...
95	Fast Food Me Gusta	Restaurant	36.454663	7.431589	0.00	0.0	5.0
96	Texas waffles	Restaurant	36.461710	7.436385	0.00	0.0	4.8
97	Sebila 1 Restaurant	Restaurant	36.464002	7.431172	0.00	0.0	3.0
98	Sebila Restaurant	Restaurant	36.442423	7.433711	0.00	0.0	NaN
99	El Bistro 24	Bar & grill	36.461873	7.435947	58.25	0.0	4.7

100 rows x 7 columns

FIGURE 2.1 – Un exemple d’une partie de la collection des données qui concerne les restaurants.

2.2.2 Prétraitement :

Nous avons fait un traitement léger sur notre ensemble de données, en supprimant les lieux qui n’ont pas de commentaires ou d’évaluations et aussi tous les lieux en dehors de la wilaya de Guelma.

2.2.3 Analyse des sentiments :

Le vecteur de caractéristique de chaque lieu comme il est présenté dans la figure précédente contient le nom et la catégorie, ainsi que la position (latitude et longitude). Ensuite, pour tous les commentaires des utilisateurs, nous avons calculé le SCORE

VADER positif et négatif.

- VADER nous donne un score combiné qui indique si la phrase est positive, ou négative [26].

Score positif :

Nous effectuons le calcul de la somme des scores positifs de chaque commentaire à partir du commentaire i jusqu'au commentaire N (le nombre total de commentaires), et ensuite on divise le résultat par N .

$$\sum_{i=1}^N (\text{valeur positive pour le commentaire})_{(i)} / N [6]$$

Un exemple réel depuis la base :

Après avoir calculé le score total des commentaires des utilisateurs pour le lieu "Café la princesse" en utilisant la méthode VADER, nous avons constaté que le score est positif (Lieu avec un score positif de 65%).

23	Café La Princesse	Grocery store	36.4643	7.42558	65.0	0.0	5.0
----	-------------------	---------------	---------	---------	------	-----	-----

FIGURE 2.2 – Un exemple réel positif depuis la base.

Score négatif :

Nous effectuons le calcul de la somme des notes négatives de chaque commentaire, à partir du commentaire i jusqu'à N (le nombre total de commentaires), et ensuite on divise le résultat par N .

$$(i = 1)^N (\text{valeur négative pour le commentaire})_{(i)} / N [7]$$

Un exemple réel depuis la base :

En calculant le score total des avis des utilisateurs sur le restaurant "Le bon coin Food" à l'aide de la méthode VADER, nous trouvons que le score est négatif (Lieu avec négativité 14.6%).

06	Le Bon Coin Food	Restaurant	36.4647	7.42853	0.0	14.6	2.0
----	------------------	------------	---------	---------	-----	------	-----

FIGURE 2.3 – Un exemple réel négatif depuis la base.

2.2.4 Importation de la Base de données :

À ce stade, après avoir effectué le traitement nécessaire, l'ensemble de données est importé vers une base de données MySQL. Le schéma ci-dessous (figure 4) représente la conception de notre base de donnée :

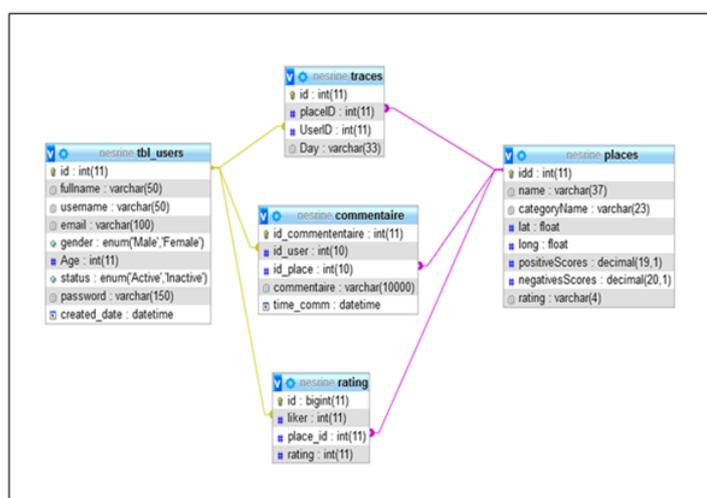


FIGURE 2.4 – conception de la base de données.

2.3 Fonctionnement du système de recommandation

Les informations sur les lieux, telles que la position et le score, sont importées dans la base de données d'un site web que nous avons réalisé appelé « FindToMe ».

Le site va recommander des lieux intéressants aux utilisateurs selon leur position et la distance des lieux ainsi que leurs notes d'évaluation (ratings et commentaires) préalablement attribuées par les utilisateurs.

Par exemple, si un nouvel utilisateur est à la recherche d'un restaurant, le système va le lui recommander en fonction de la distance, l'évaluation du lieu et la positivité du score VADER.

Ensuite, nous demandons à ce dernier de noter ou de faire un commentaire sur ce lieu et notre système calcule le score VADER de son commentaire. S'il est positif, nous recommandons ce restaurant encore une fois, autrement le système ne recommandera jamais ce lieu et il sera ajouté à la liste des lieux noirs pour cet utilisateur.

Pour améliorer la précision de la recommandation des lieux faite par notre algorithme, nous le basons sur la collecte des traces des utilisateurs inscrits sur notre site. Le système va apprendre le comportement des utilisateurs en fonction de leurs recherches, de leurs notes, et de leurs commentaires. Par exemple, lorsqu'un utilisateur

visitera tous les dimanches le même lieu, le système va le recommander comme étant un lieu intéressant (qui a un excellent score VADER ou une excellente note) pour le dimanche prochain.

Le schéma suivant résume le fonctionnement du système « FindToMe » :

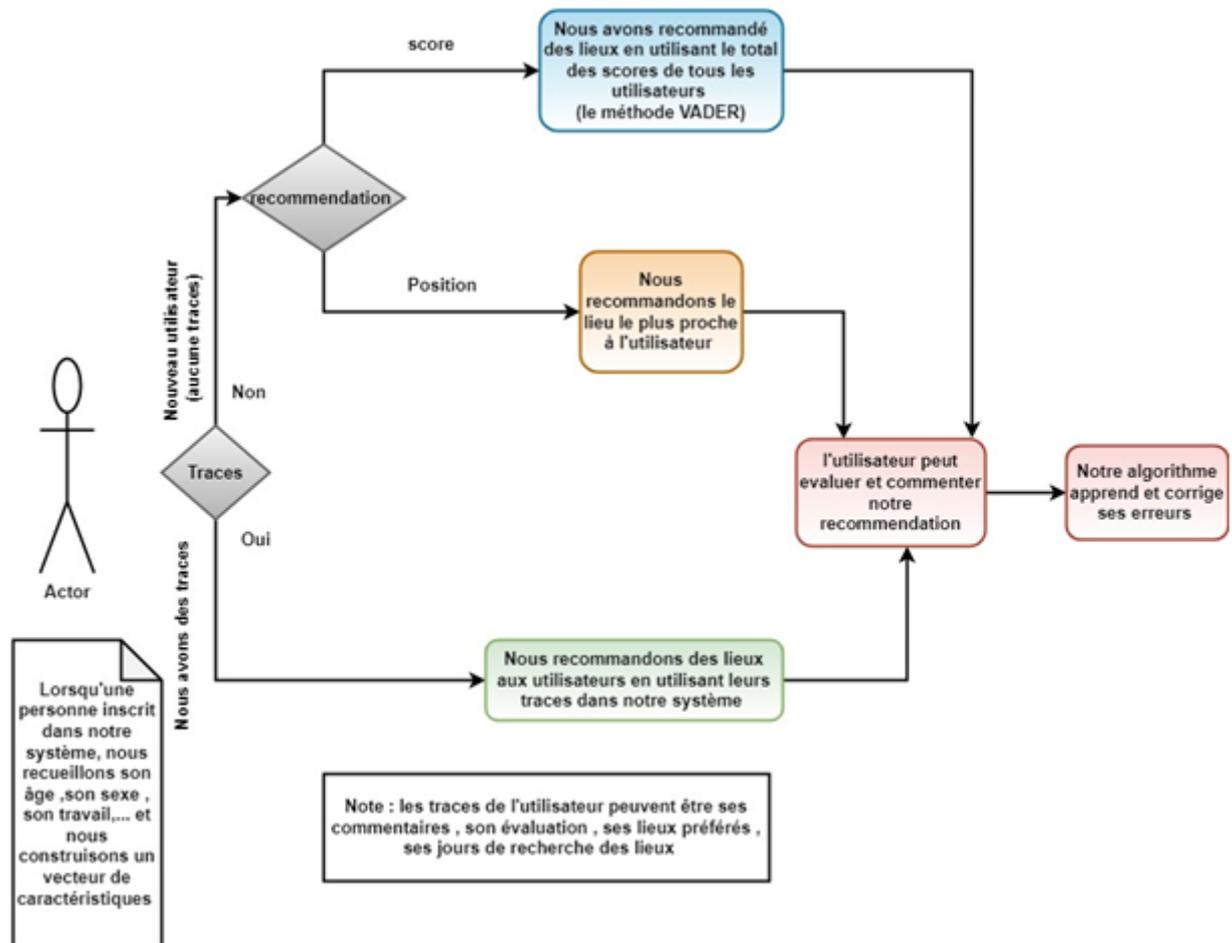


FIGURE 2.5 – Fonctionnement du système FindToMe

2.4 Conception du site FindToMe :

2.4.1 Diagramme de cas d'utilisation « FindToMe »

Le diagramme de cas d'utilisation est un diagramme UML utilisé pour donner une vision globale du comportement fonctionnel d'un système logiciel. Un cas d'utilisation représente une unité discrète d'interaction entre un utilisateur (Humain ou Machine) et un système. Dans un diagramme de cas d'utilisation il existe des acteurs qui interagissent avec des cas d'utilisation UC.

Les cas d'utilisation permettent de structurer les besoins des utilisateurs et les objectifs du système [16]. La figure 6 ci-dessous représente le diagramme de cas d'utilisation de notre système FindToMe.

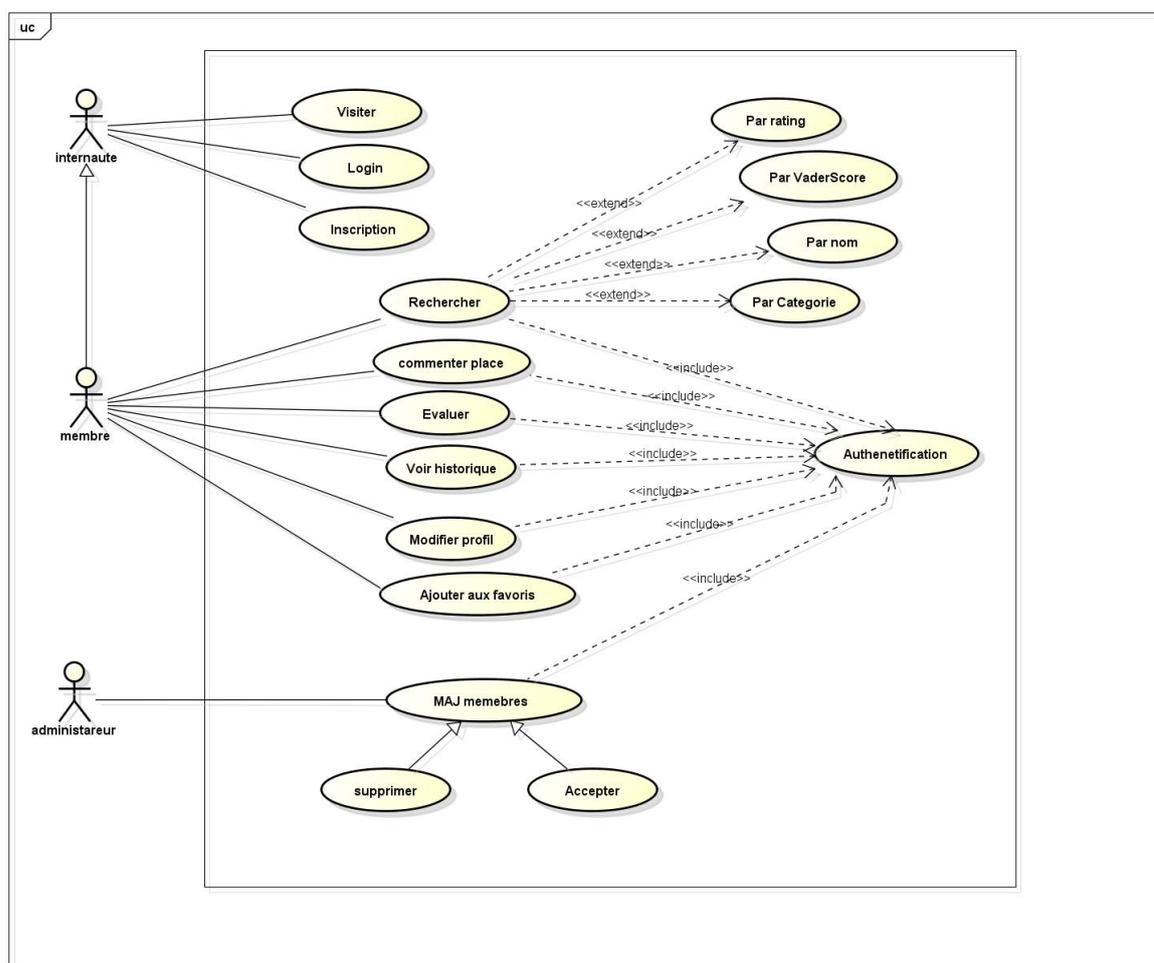


FIGURE 2.6 – Diagramme de cas d’utilisation « FindToMe »

L’administrateur est le responsable de notre système, il fait la mise à jour en supprimant et acceptant les demandes d’inscription.

l’ensemble des internautes peuvent visiter notre site mais une fois leurs inscriptions sont validées ils peuvent rechercher des places, commenter, évaluer, voir l’historique, modifier profil, et ajouter aux favoris.

2.4.2 Diagramme de classe « FindToMe » :

Les diagrammes de classe décrivent clairement la structure d’un système particulier en modélisant ses classes, ses attributs, ses opérations et les relations entre ses

objets [25].

La figure 7 ci-dessous représente le diagramme de classe de notre système FindToMe.

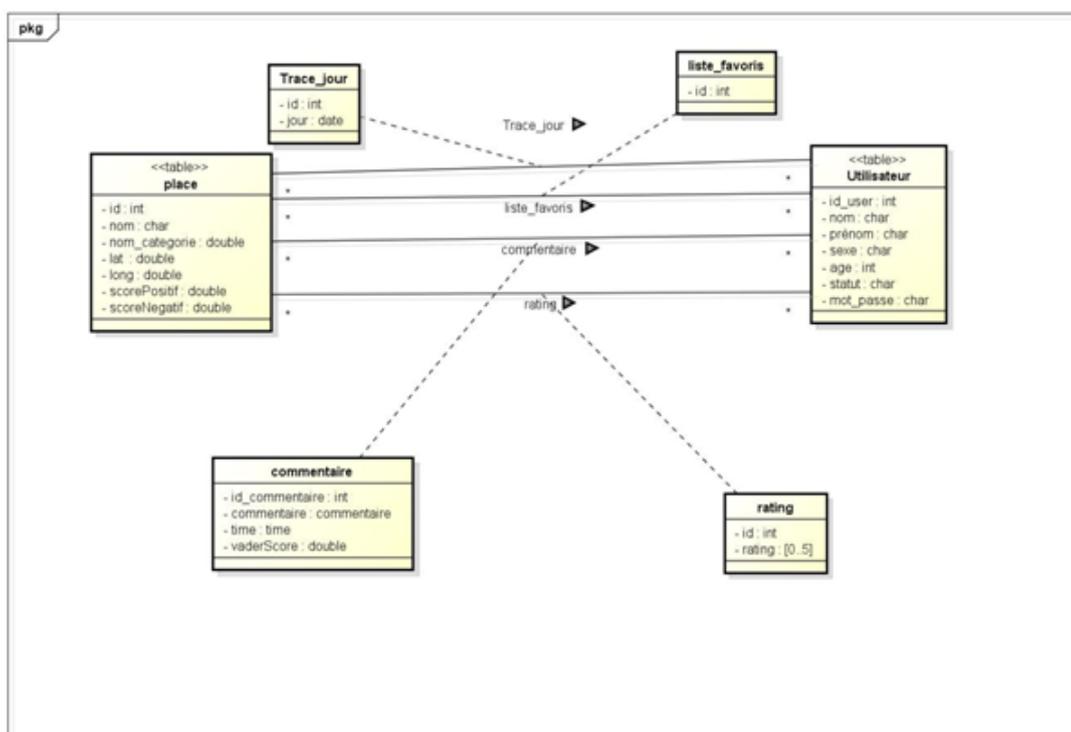


FIGURE 2.7 – Diagramme de classe « FindToMe »

2.4.3 Diagramme de séquence :

Un diagramme de séquence représente la séquence de messages entre les objets au cours d'une interaction. Un diagramme de séquence comprend un groupe d'objets, représentés par des lignes de vie, et les messages que ces objets échangent lors de l'interaction [17].

Les figures 8, 9 et 10 ci-dessous représentent trois interactions de l'utilisateur avec notre système FindToMe.

Diagramme de séquence « Login »

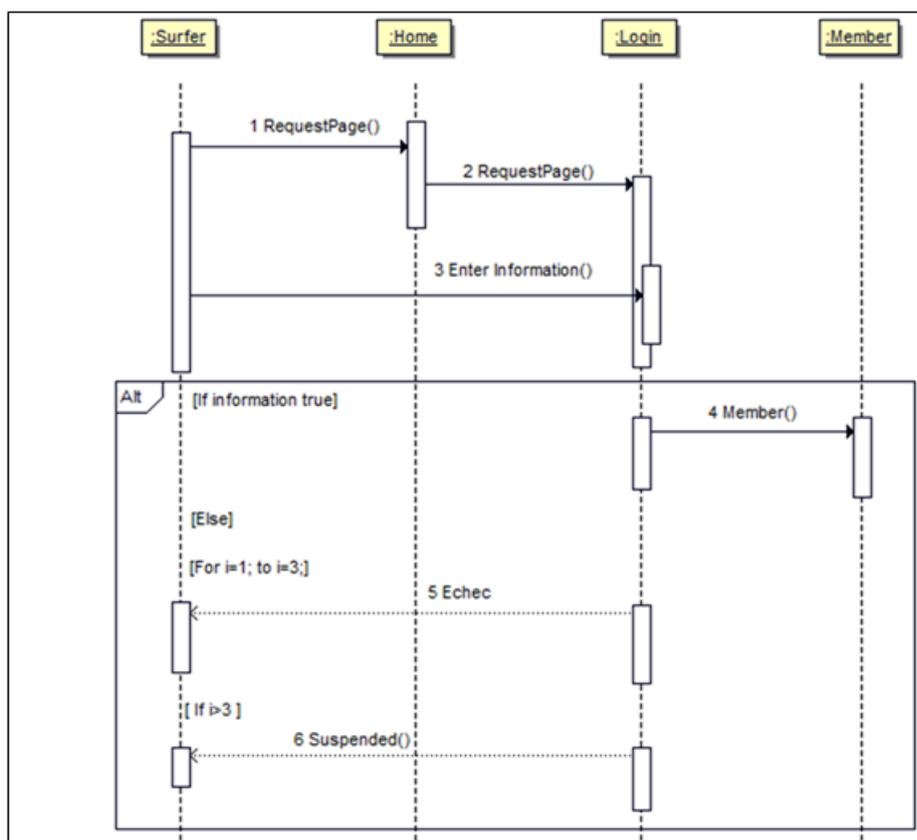


FIGURE 2.8 – Diagramme de séquence « Login »

Chaque membre doit remplir le formulaire de connexion avec son nom d'utilisateur et son mot de passe. Le système vérifie l'existence de ce compte dans la base de données, s'il existe alors l'authentification est réussite sinon l'utilisateur doit refaire l'opération.

En cas d'erreur soit dans le nom de l'utilisateur ou du mot de passe le membre a le droit de refaire l'opération trois (03) fois sinon il sera suspendu de notre site.

Diagramme de séquence « Inscription »

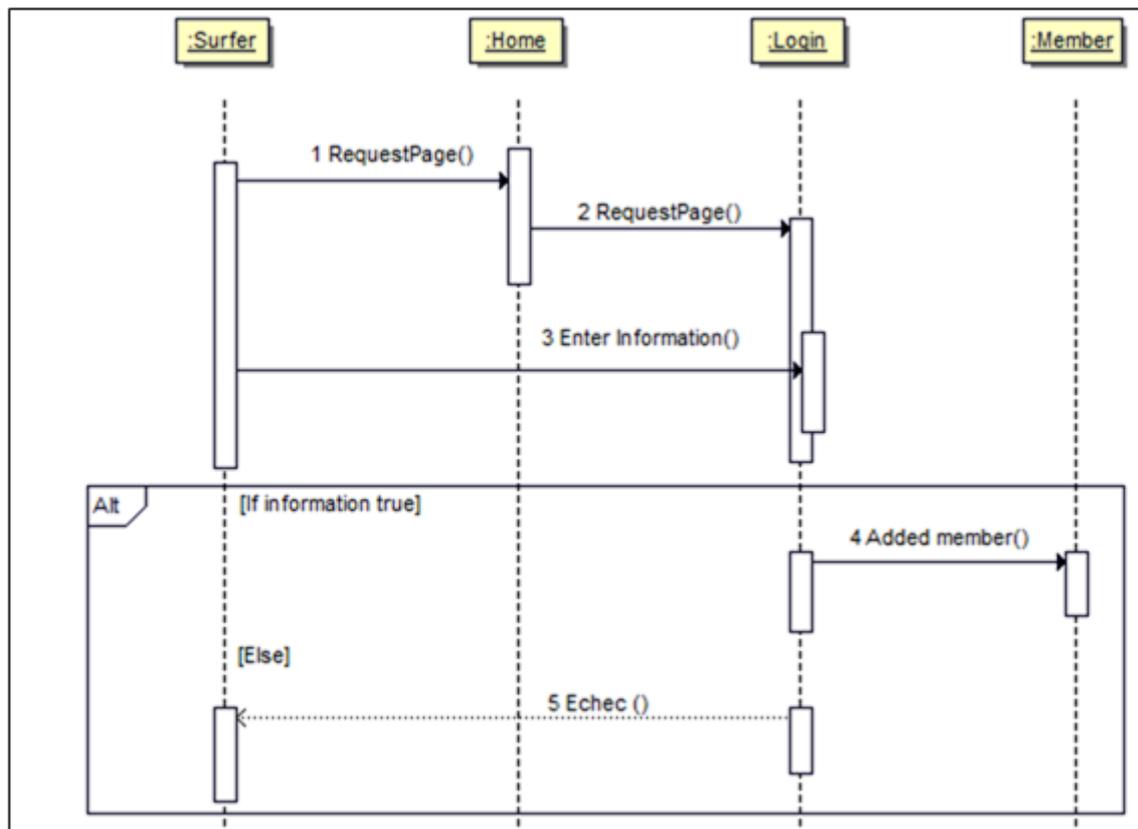


FIGURE 2.9 – Diagramme de séquence « Inscription »

Pour s'inscrire, l'internaute doit remplir le formulaire d'inscription qui contient : le nom, le nom d'utilisateur, l'email, l'âge et le sexe.

Le système fait une vérification, si les informations saisies sont correctes, alors l'internaute sera ajouté à la liste des membres sinon le système lui renvoie un message d'échec.

Diagramme de séquence « ajouter au favori »

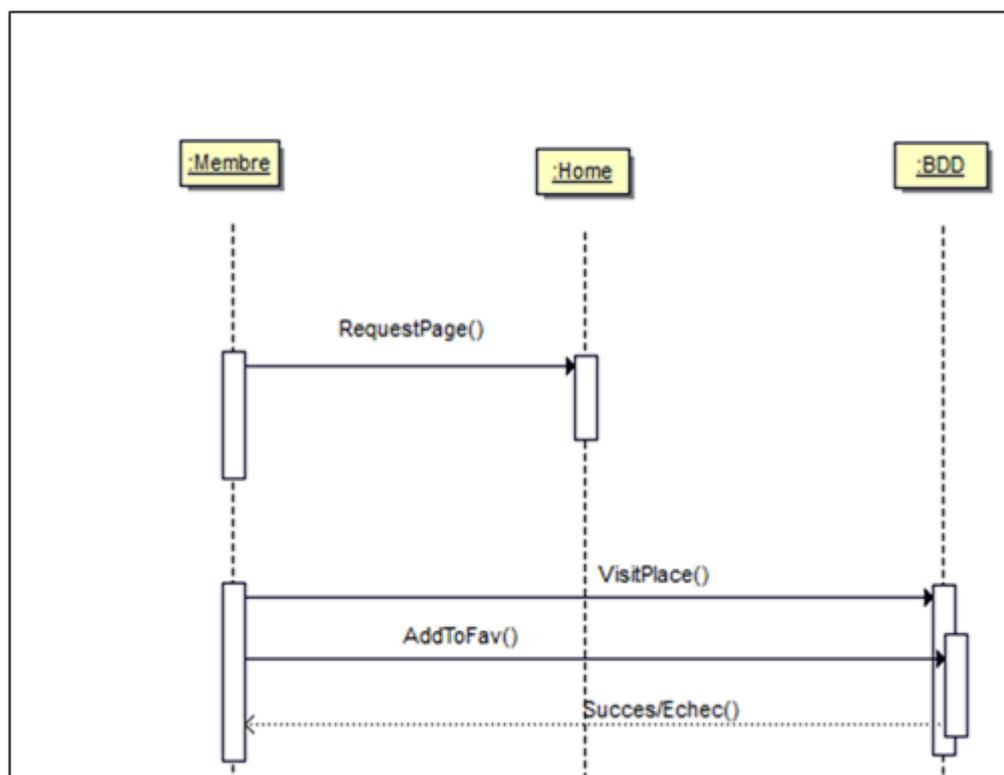


FIGURE 2.10 – Diagramme de séquence « ajouter aux favoris »

Après l'authentification et la visite sur place, si l'utilisateur est satisfait il pourra l'ajouter à sa liste des favoris pour que le système lui recommandera une autre fois.

2.5 Algorithme de recommandation

L'algorithme de recommandation des lieux est présenté comme suit :

Algorithme FindToMe :

Entrée : L'utilisateur cherche le lieu en utilisant la catégorie (par exemple, l'utilisateur voudrait aller chez un médecin).

Sortie : Les 3 médecins les plus adaptés en fonction de notre système de recommandation.

Début :

Il y a deux scénarios possibles dans notre algorithme

Cas 1 : Utilisateur qui ne possède aucune trace (démarrage à froid)**La sélection de l'option de recherche :**

1. Le lieu le plus proche de l'utilisateur :

Le système calcule la distance entre tous les résultats possibles, puis il choisit les 3 lieux les plus proches de la position actuelle de l'utilisateur. Par la formule de distance euclidienne suivante :

$$D(p, q) = \sqrt{\sum_{i=0}^n (q_i - p_i)^2} [8]$$

2. La note la plus élevée (note >4) :

Le système de recommandation effectue un tri des lieux en fonction des notes d'évaluation (rating) données par les gens qui ont visité ces lieux, et propose trois lieux qui ont la plus haute note. Si plusieurs lieux possèdent la même évaluation, il en sélectionnera 3 de façon aléatoire.

3. Analyse des sentiments (méthode VADER) :

Le système de recommandation effectue un calcul de score pour tous les commentaires qui concerne le lieu ou bien la catégorie sélectionnée par l'utilisateur par exemple (les médecins) et renvoie trois lieux ayant le score le plus élevé tel que : un score positif est strictement supérieur à 0.5 alors le système sélectionne les lieux ayant plus de commentaires positifs.

Cas 2 : utilisateur avec des traces :

Le système recommandera à l'utilisateur 2 lieux (selon la méthode VADER ou selon le score de son évaluation) à partir de la liste de ses favoris, autrement dit :

1. Les deux lieux les mieux notés par l'utilisateur lui-même.
2. Les deux lieux qui ont obtenu le score le plus positif selon la méthode VADER.

- Le système recommande toujours les lieux les plus visités par l'utilisateur depuis la table des traces.
- Si l'utilisateur est un enfant moins de 13 ans, le système effectue un filtrage de localisation (par exemple, Les enfants ne cherche pas à trouver des avocats).

Fin.

2.6 Conclusion

L'étape conceptuelle est une étape essentielle pour la réussite de tout projet. Elle facilite la réalisation et la mise en œuvre du projet. Cela nous a permis de répondre aux questions sur la mise en place du système que nous étions en train de développer, ainsi que de trouver les moyens et les outils dont nous avons besoin pour concrétiser notre système et traiter le jeu de données que nous avons collecté.

CHAPITRE 3

CHAPITRE 03 : IMPLÉMENTATION

3.1 Introduction

Dans ce chapitre on entame la réalisation du système précédemment conçu. On commence par présenter l'environnement de développement et les bibliothèques utilisées. Ensuite, on présente les outils et les langages utilisés dans l'implémentation. Puis, on aborde en détail les étapes de travail, et on termine par la discussion des résultats obtenus.

3.2 Environnements de développement et langages utilisés

L'environnement de développement regroupe l'ensemble des outils de gestion informatiques qui participent à la conception et la création de notre site web. De nos jours beaucoup d'outils d'implémentation des sites web sont mis à la disposition des utilisateurs, on va donc citer ceux qu'on a utilisés. On cite également les langages de programmation utilisés afin de réaliser ce projet.

3.2.1 PHP (Hypertext Preprocessor) :

Le PHP a été choisi comme étant un langage de programmation libre, spécialement adopté pour créer des pages web actives via un serveur HTTP et ainsi agir comme n'importe quel langage déchiffré de façon locale. Aussi, PHP permet le codage absolu orienté objet [38].

3.2.2 Sublime text editor :

Sublime text se définit comme un « éditeur de texte sophistiqué pour le code, le balisage et la prose » initialement, en janvier 2008, on est maintenant à sa quatrième version majeure. Il prend en charge Mac, Windows et Linux. Sublime Text ressemble essentiellement à un couteau suisse qui peut être appliqué à n'importe quel cas d'utilisation ou difficulté impliquant la manipulation de texte. Contrairement aux autres éditeurs de texte, Sublime Text permet d'exécuter le code directement dans l'application. Ceci spécialement efficace pour les projets simples comme les petits scripts Bash ou Python [45].

3.2.3 HTML (Hypertext Markup Language), CSS (Cascade StyleSheet) :

HTML et CSS décrivent comment une page web est présentée. HTML est l'ossature du web. Il arrange les différents éléments d'une page web ainsi le CSS permet de donner un style visuel pour ces éléments [23].

3.2.4 Leaflet :

Leaflet est une bibliothèque fondamentale JavaScript à code source ouvert pour les cartes interactives adaptées aux téléphones portables. Leaflet ne dépasse pas les 42 Ko de JS, elle offre toutes les fonctionnalités aux quelles ont recours certains développeurs.

Leaflet avait été élaboré avec soin et limpidité, réussite et convivialité. Elle agit rentablement sur les différentes plates-formes de bureaux ainsi que les mobiles, peut être amplifiée avec un nombre abondant de plugins, détient une API somptueuse, agréable à utiliser et bien documentée, et un code source accessible et déchiffrable auquel il est attirant d'y prendre part [31].

3.2.5 Apify :

Est un robot web hébergé qui favorise à toute personne qui dispose des capacités et des connaissances essentielles en programmation d'extraire des données organisées de n'importe quel site web.

A l'opposé du Scraping web par pointer-cliquer, apify agit sur de nouveaux sites web bien plus actifs et compliqués.

On a eu recours à apify afin d'acquérir les différentes informations concernant les endroits situés à Guelma à partir de Google Maps [9].

3.2.6 JavaScript :

JavaScript (souvent abrégé en « JS ») est un langage de script léger, orienté objet, généralement connu comme le langage de script des pages, utilisé pour rendre les pages interactives. JS permet d'implémenter des mécanismes complexes sur une page web.

À chaque fois qu'une page web fait plus que simplement afficher du contenu statique afficher du contenu mis à jour à des temps déterminés, des cartes interactives, des animations 2D/3D, des menus vidéos défilants, etc.

C'est la troisième couche des technologies standards du web, les deux premières sont HTML et CSS. Pour rendre nos pages interactives nous avons utilisé le langage JavaScript et sa bibliothèque JQuery [27].

3.2.7 MySQL :

Un système de gestion de bases de données relationnelles (SGBDR). Il appartient aux logiciels de gestion de base de données les plus utilisés au monde compte tenu de sa simplicité [34].

3.2.8 Bootstrap :

Bootstrap est une infrastructure de développement frontale, gratuite et open source pour l'innovation des sites et d'applications Web. Bootstrap repose sur HTML, CSS et JavaScript (JS) pour rendre le développement des sites et d'applications réactives et tout-mobile plus facile [11].

3.2.9 AJAX :

Le terme AJAX désigne une technologie popularisée dans le domaine de la création des sites internet. Elle est utilisée pour apporter de l'interactivité aux pages d'un site web tout en économisant les ressources serveur.

Il permet de communiquer avec le serveur à l'aide de code Javascript en arrière-plan pendant que la page est affichée à l'écran. Ainsi le contenu de la page peut être modifié sans qu'il soit nécessaire de faire transiter et afficher la page en entier.

Ajax est utilisée pour la mise à jour des formulaires et des paniers sur la plupart des sites web. C'est une technologie asynchrone : le code de la page continue de s'exécuter pendant que l'appel vers le serveur est effectué [6].

3.2.10 GOOGLECOLAB :

C'est un produit de Google Research il permet aux utilisateurs d'écrire et d'exécuter le code Python de son choix par le biais du navigateur.

C'est un environnement adapté à l'apprentissage automatique (machine learning), à l'analyse de données (data analysis) et à l'éducation (e-learning).

Nous pouvons dire que GoogleColab est un service hébergé de notebooks Jupyter, il ne nécessite aucune configuration et permet d'accéder à des ressources informatiques gratuitement dont des GPU (Graphics Processing Unit) [22].

3.2.11 Python :

Python est le langage de programmation le plus connu entre les informaticiens ces dernières années. Il est propulsé en tête de la gestion d'infrastructure, d'analyse de données ou dans le domaine du développement de logiciels. Python permet aux développeurs de se concentrer sur ce qu'ils font plutôt que sur la manière dont ils le font.

Il a libéré les développeurs des contraintes de formes qui occupaient leur temps avec les langages plus anciens. Nous pouvons dire que développer du code avec Python est beaucoup meilleure qu'avec d'autres langages [40].

3.2.12 NumPy :

C'est le package fondamental pour le calcul scientifique en Python.

Il s'agit d'une bibliothèque Python qui fournit un objet tableau multidimensionnel, différents objets dérivés comme les tableaux et les matrices masqués et un assortiment de routines pour des opérations rapides sur des tableaux, spécialement mathématiques, logiques, manipulation de forme, tri, sélection, nous avons utilisé cette bibliothèque pour adapter le type d'entrée et la manipulation des tableaux [35].

3.2.13 Pandas :

Pandas DataFrame est une structure de données tabulaires bidimensionnelle à taille variable, éventuellement hétérogène, avec des axes étiquetés (lignes et colonnes). Un DataFrame est une structure de données bidimensionnelle, dont les données sont alignées de manière tabulaire en lignes et en colonnes. Le DataFrame de Pandas se compose de trois éléments principaux, les données, les lignes et les colonnes [36].

3.2.14 VADER (Valence Aware Dictionary for Sentiment Reasoning) :

VADER est un model utilisé pour l'analyse des sentiments dans les textes, sensible à la fois à la polarité (positive/négative) et à l'intensité (force) des émotions. disponible dans le package NLTK et peut être appliqué directement à des données textuelles non étiquetées.

L'analyse sentimentale VADER s'appuie sur un dictionnaire qui associe des caractéristiques lexicales à des intensités d'émotions connues sous le nom de score sentimental. Le score de sentiment d'un texte peut être obtenu en additionnant l'intensité de chaque mot du texte [50].

3.2.15 JQuery :

JQuery est une bibliothèque JavaScript gratuite, libre et multiplateforme compatible avec l'ensemble des navigateurs Web (Internet Explorer, Safari, Chrome, Firefox, etc.), elle a été conçue et développée en 2006 pour faciliter l'écriture de scripts.

Il s'agit du Framework JavaScript le plus connu et le plus utilisé.

Il permet d'agir sur les codes HTML, CSS, JavaScript et AJAX et s'exécute essentiellement côté client. La bibliothèque JQuery permet de gagner en rapidité dans l'interaction avec le code HTML d'une page Web [28].

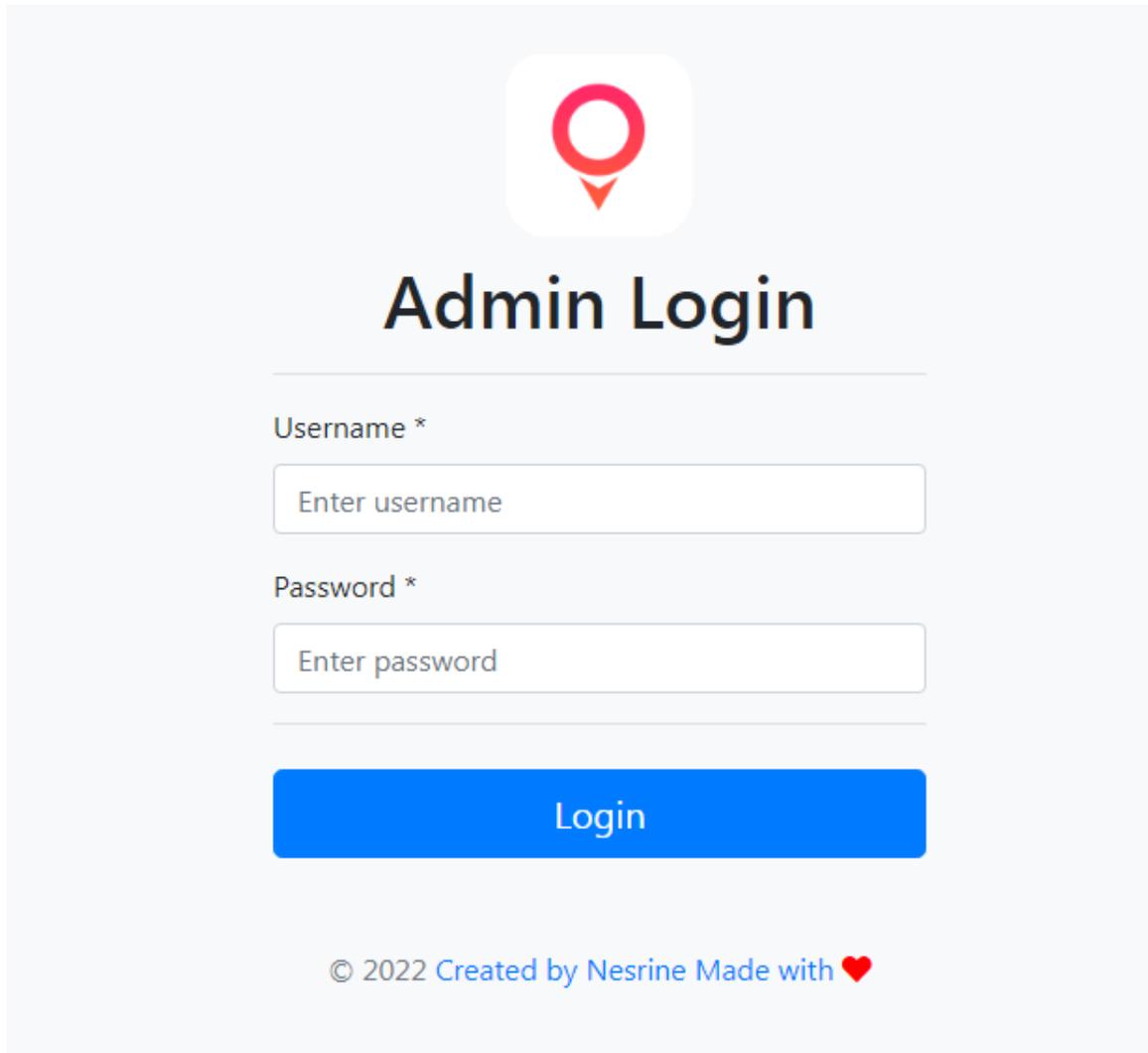
3.2.16 AppsGeyser :

AppsGeyser est une plateforme d'outils en ligne pour les applications Android. Fournit des outils importants que les créateurs d'applications android. Cest à dire que tout le monde peut utiliser les outils de création d'applications d'AppsGeyser, qu'il s'agisse d'un développeur expérimenté ou de quelqu'un qui n'a pas de connaissances techniques.

Aujourd'hui, plus d'un demi-million d'applications Android ont été créées à l'aide des outils d'AppsGeyser et plus de 350 millions d'applications ont été téléchargées. Nous avons utilisé APPGEYSER pour rendre notre site web une application mobile [10].

3.3 Les interfaces graphiques du système (Findtome) :

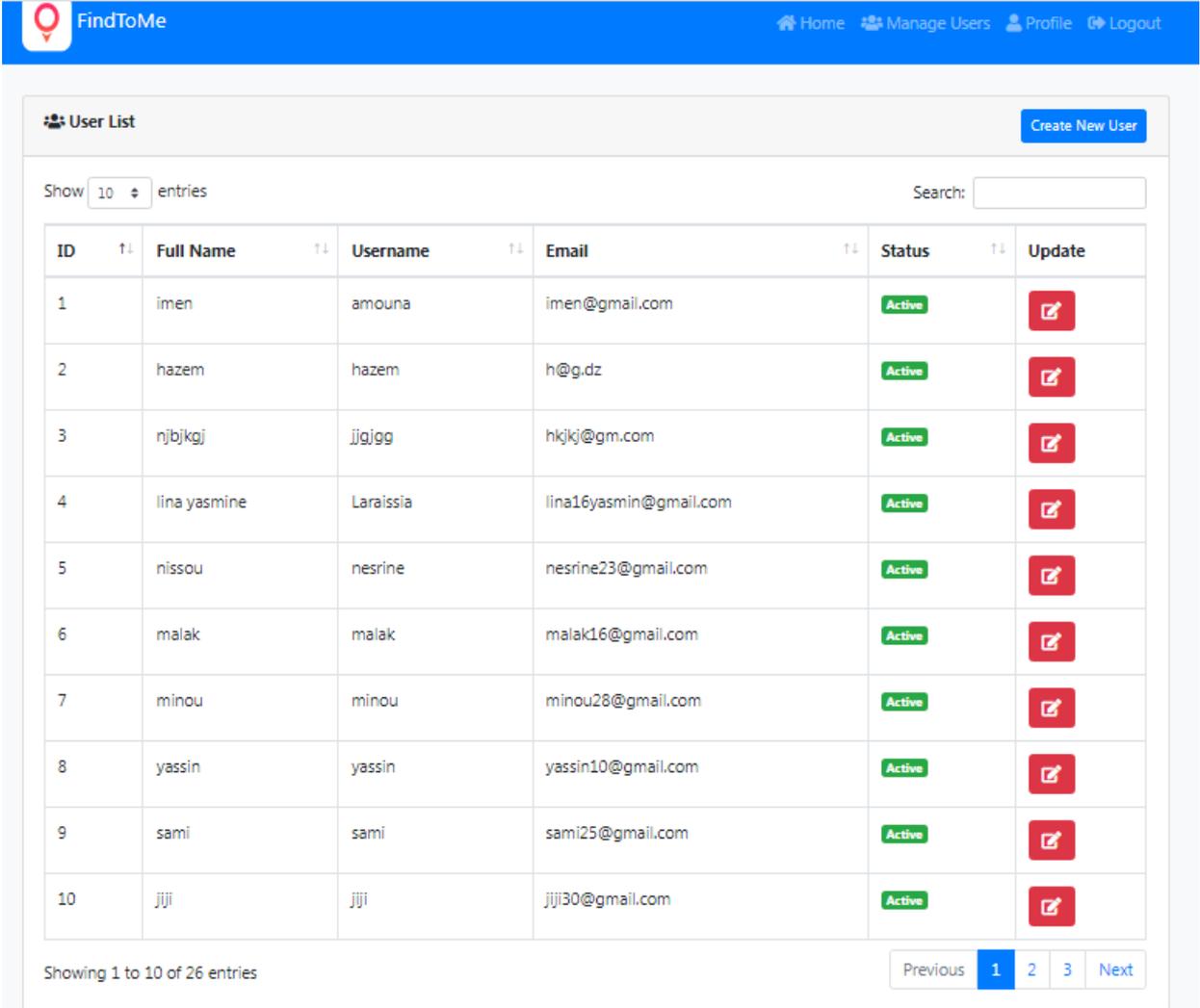
3.3.1 Espace Administrateur :



The image shows a login interface for an administrator. At the top center is a red location pin icon inside a white rounded square. Below this is the text "Admin Login" in a large, bold, black font. Underneath, there are two input fields: "Username *" and "Password *". Each field has a light gray border and contains the placeholder text "Enter username" and "Enter password" respectively. Below the input fields is a blue button with the text "Login" in white. At the bottom of the interface, there is a copyright notice: "© 2022 Created by Nesrine Made with ❤️".

FIGURE 3.1 – Espace administrateur

L'administrateur est le responsable du site Findtome, il s'authentifie via un nom d'utilisateur et un mot de passe.



The screenshot displays the 'User List' interface for the 'FindToMe' application. At the top, there is a navigation bar with the application logo and links for Home, Manage Users, Profile, and Logout. Below the navigation bar, the 'User List' section includes a 'Create New User' button and a search bar. The main content is a table with 10 rows of user data. Each row contains an ID, Full Name, Username, Email, Status (all 'Active'), and an 'Update' button. The table is paginated, showing 1 to 10 of 26 entries.

ID	Full Name	Username	Email	Status	Update
1	imen	amouna	imen@gmail.com	Active	
2	hazem	hazem	h@g.dz	Active	
3	njbkgj	jjgg	hkjkj@gm.com	Active	
4	lina yasmine	Laraissia	lina16yasmin@gmail.com	Active	
5	nissou	nesrine	nesrine23@gmail.com	Active	
6	malak	malak	malak16@gmail.com	Active	
7	minou	minou	minou28@gmail.com	Active	
8	yassin	yassin	yassin10@gmail.com	Active	
9	sami	sami	sami25@gmail.com	Active	
10	jji	jji	jji30@gmail.com	Active	

FIGURE 3.2 – la liste des utilisateurs du système "findtome".

Dans cette page l'administrateur établit la mise à jour des utilisateurs (acceptation, suppression, modification du profil, création des nouveaux utilisateurs).

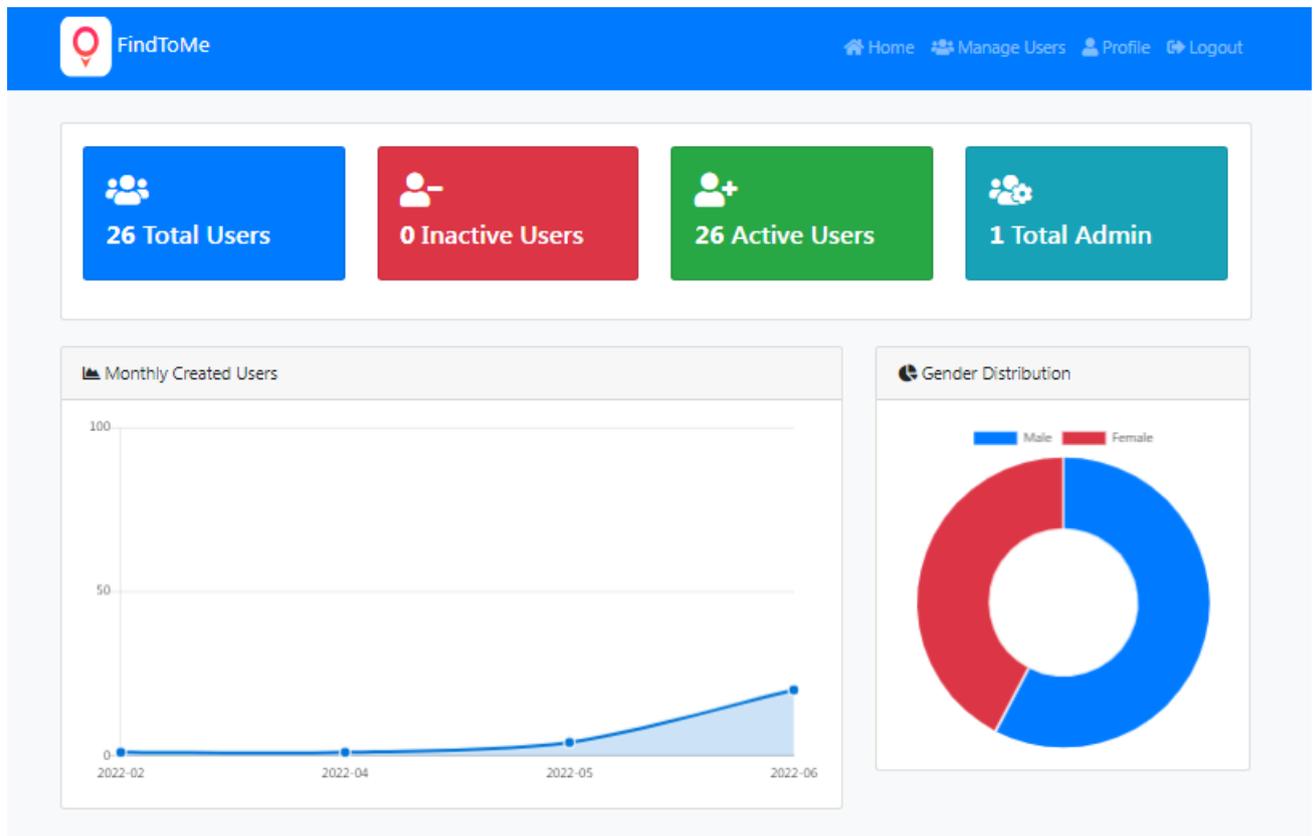


FIGURE 3.3 – l'ensemble de données présentées par le dashboard.

L'administrateur contrôle le système depuis cet espace, et procède à la visualisation des statistiques (nombre d'utilisateurs actifs, inactifs) tout en comptant le nombre d'hommes et de femmes qui utilisent le système Findtome.

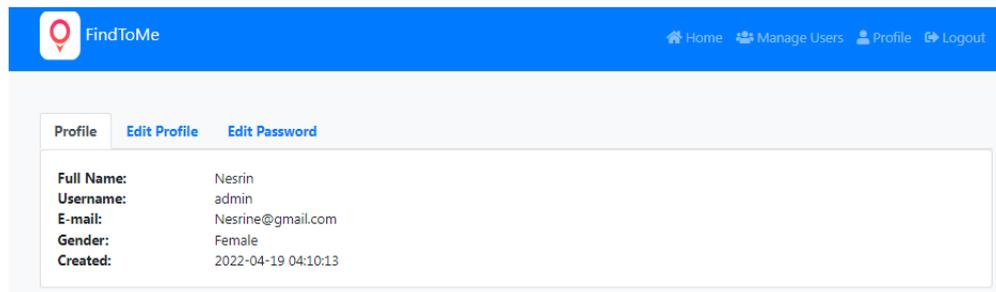


FIGURE 3.4 – Profil administrateur.

À l'aide de cette page l'administrateur peut consulter son profil et faire des modifications.

3.3.2 Espace utilisateur :

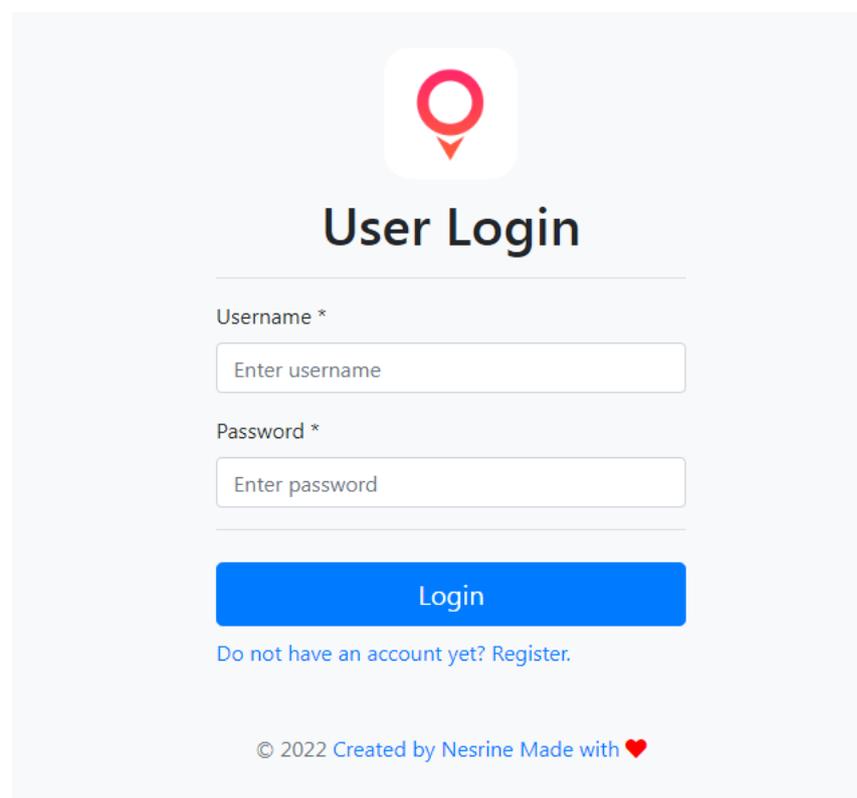
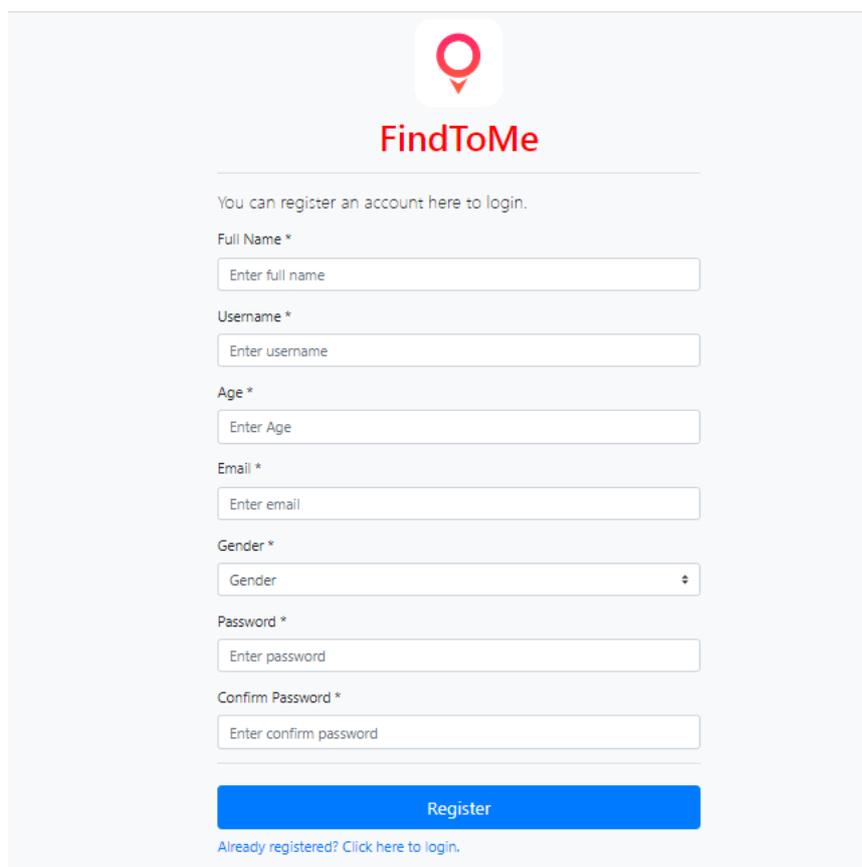


FIGURE 3.5 – Connexion au système.

Cette page permet aux visiteurs de notre plateforme de s'authentifier via un nom d'utilisateur et un mot de passe avec une option d'inscription s'ils ne possèdent pas de compte pour accéder aux différentes fonctionnalités de l'application.



The image shows a registration form for a system named "FindToMe". At the top center is a logo consisting of a red circle with a white location pin inside, and the text "FindToMe" in red below it. Below the logo, there is a horizontal line and the text "You can register an account here to login." followed by several input fields, each with an asterisk indicating it is required:

- Full Name ***: A text input field with the placeholder "Enter full name".
- Username ***: A text input field with the placeholder "Enter username".
- Age ***: A text input field with the placeholder "Enter Age".
- Email ***: A text input field with the placeholder "Enter email".
- Gender ***: A dropdown menu with "Gender" selected and a small downward arrow on the right.
- Password ***: A text input field with the placeholder "Enter password".
- Confirm Password ***: A text input field with the placeholder "Enter confirm password".

At the bottom of the form is a blue button with the text "Register". Below the button is a link: "Already registered? Click here to login."

FIGURE 3.6 – Formulaire d’inscription du système.

À l’aide de cette page, les internautes peuvent créer des comptes. Il suffit de remplir ces champs avec les informations qui les correspondent (le nom, le nom d’utilisateur, l’âge, l’email, le sexe et le mot de passe).

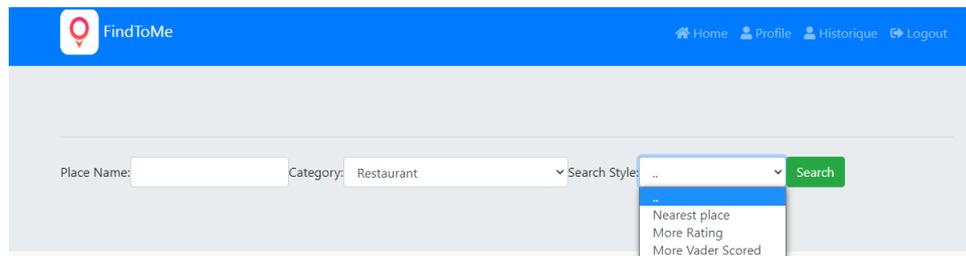


FIGURE 3.7 – Page d’accueil du système

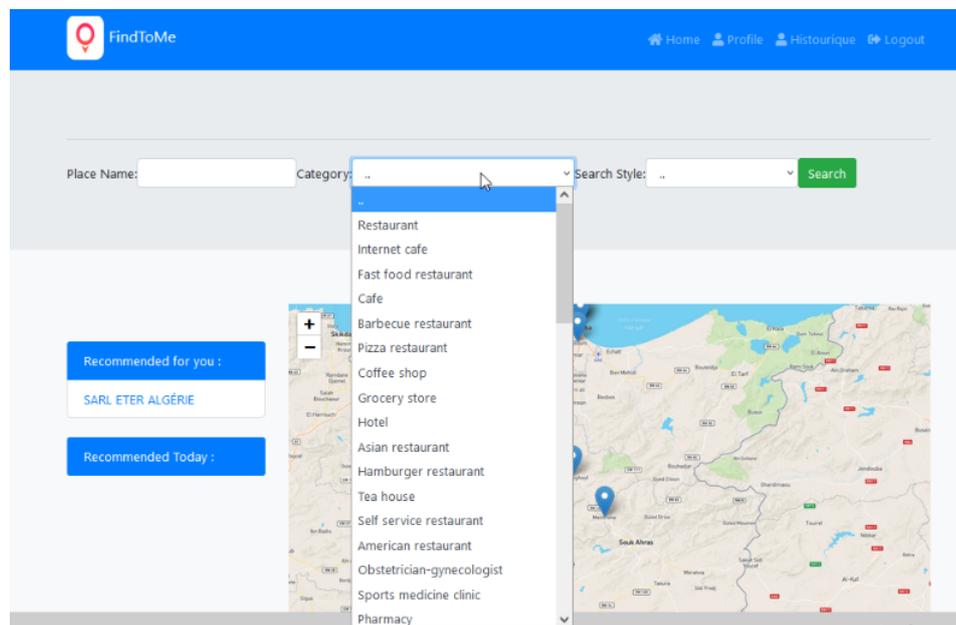
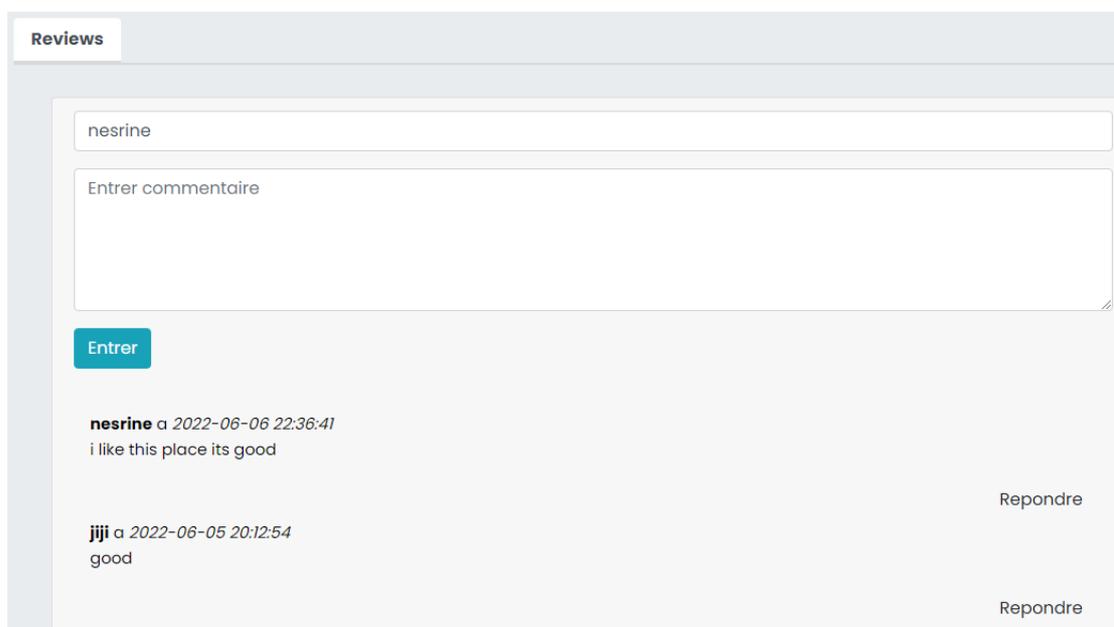


FIGURE 3.8 – l’ensemble des catégories disponibles sur le système.

Après l’authentification, et une fois l’utilisateur accède au site, trois (03) cases seront affichées :

- 1) Place name : l’utilisateur mentionne le nom du lieu souhaité, s’il veut effectuer une recherche d’un endroit bien précis.
- 2) Catégorie : l’utilisateur choisit la catégorie du lieu souhaité (plusieurs catégories lui seront affichées : restaurant, internet café, médecin, avocat, pharmacie, etc.)
- 3) Search style : trois (03) options seront affichées :
 - a) nearest place : le système lui recommande trois lieux les plus proches de sa position.
 - b) more rating : le système lui recommande les trois lieux ayant les plus hautes évaluations.

c) more vader scored : le système lui recommande les trois lieux ayant plus de commentaires favorables selon l'outil d'analyse des sentiments VADER.



Reviews

nesrine

Entrer commentaire

Entrer

nesrine a 2022-06-06 22:36:41
i like this place its good

jji a 2022-06-05 20:12:54
good

Repondre

Repondre

FIGURE 3.9 – le champ dédié aux commentaires des différents utilisateurs.



FIGURE 3.10 – la rubrique dédiée aux notations par étoiles.

Après la visite du lieu recommandé par le système et choisit par l'utilisateur. Ce dernier peut donner son avis soit sous forme de commentaire ou bien par des étoiles (Rating), À l'aide de ces deux pages comme il peut lire les commentaires et consulter les évaluations des autres utilisateurs sur ces lieux.



FIGURE 3.11 – Espace d’historique des utilisateurs.

Sur cette page, l’utilisateur peut visualiser son historique (places évaluées par étoile, places évaluées par commentaires, et ses places favorites).

Exemple de recommandation d’une cafétéria :

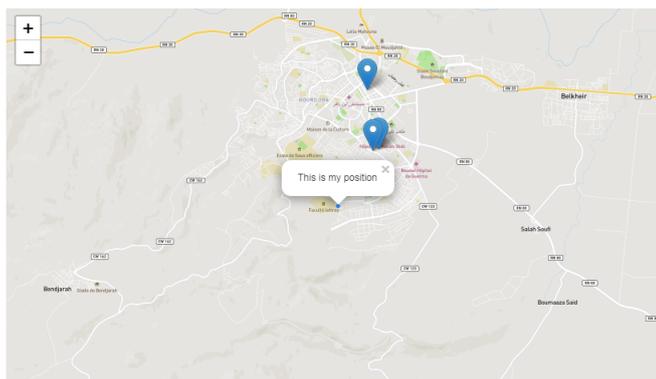


FIGURE 3.12 – Recommandation du système.

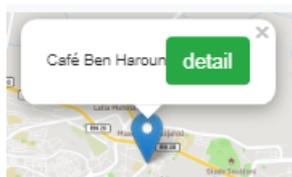


FIGURE 3.13 – Le choix de lieu.

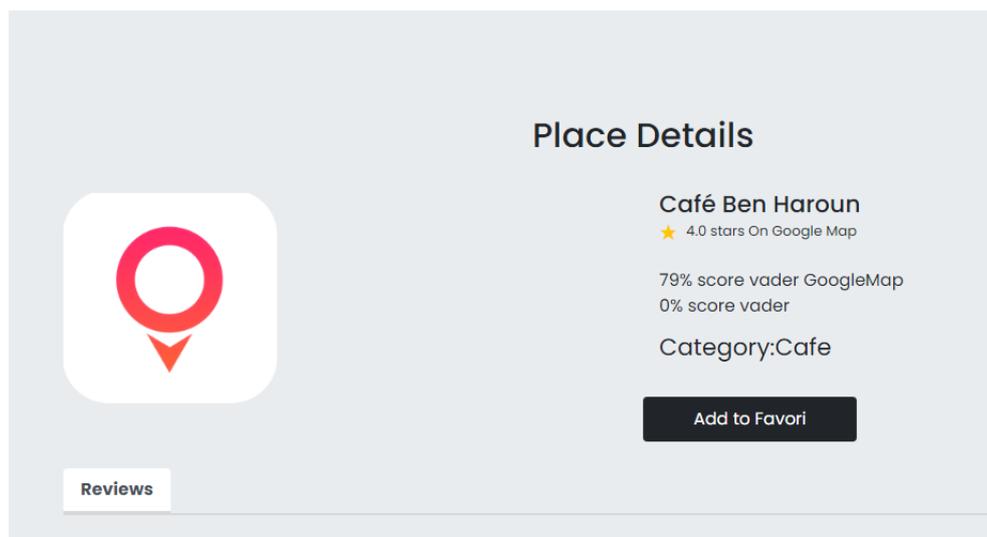
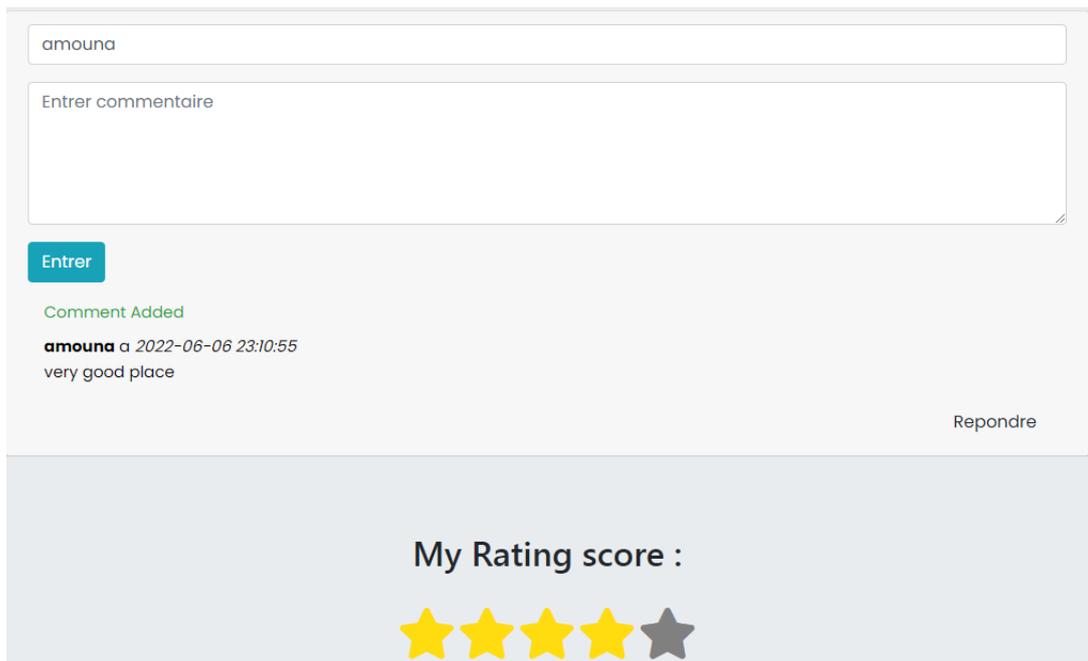


FIGURE 3.14 – Détails de lieu.

Ces captures montrent que l'utilisateur a choisi le café Ben haroun. Puis, il peut cliquer sur détails afin d'afficher les informations d'évaluation suivantes : (rating et valeur de score VADER depuis googlemap, option d'ajouter au favoris).

Nous remarquons que VADER score pour cette utilisateur est 0%, car il n'a pas encore évalué cet endroit.



The image shows a user interface for submitting a comment and displaying a rating. At the top, there is a text input field containing the name 'amouna'. Below it is a larger text area with the placeholder 'Entrer commentaire'. A blue button labeled 'Entrer' is positioned below the text area. Underneath the button, a green message reads 'Comment Added'. Below that, the comment text is displayed: 'amouna a 2022-06-06 23:10:55' followed by 'very good place'. A 'Repondre' button is located at the bottom right of the comment section. Below the comment section, the text 'My Rating score :' is displayed, followed by a row of five stars. The first four stars are yellow, and the fifth star is grey, indicating a 4-star rating.

FIGURE 3.15 – Commentaire positif.

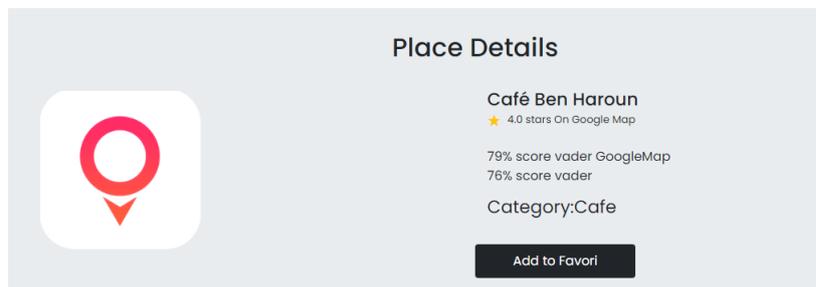


FIGURE 3.16 – le résultat obtenu après le commentaire.

Après avoir visité cet endroit et après qu'il a exprimé sa satisfaction, l'utilisateur a commenté en attribuant à ce café quatre étoiles et un commentaire très favorable. Notre système a calculé ensuite le score VADER du commentaire et lui a attribué un score positif de 76%.

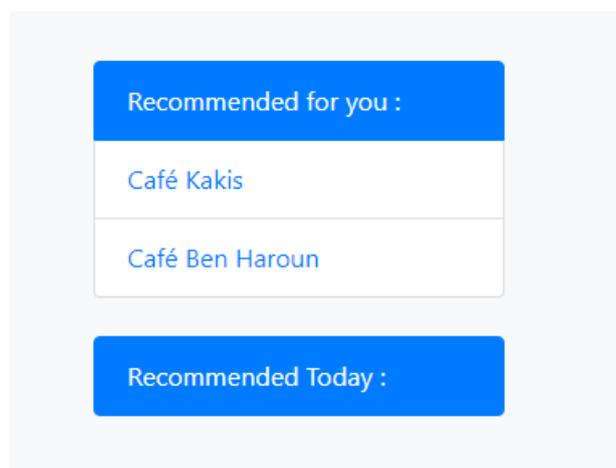


FIGURE 3.17 – Recommandation automatique.

La cafétéria est ajoutée à la liste des places qui seront recommandées à l'utilisateur par le système.

3.4 Résultat et discussion :

Dans le but d'évaluer la qualité des recommandations, on a appelé 21 utilisateurs à créer des comptes sur FindToMe et l'essayer en temps réel au niveau de la ville de Guelma.

Les utilisateurs ont essayé le système à différents endroits de la ville, et différentes recommandations ont été suggérées. Pour chaque utilisateur trois (03) lieux ont été recommandés, l'utilisateur évalue chaque lieu en attribuant un nombre d'étoiles (1-5) pour exprimer son degré de satisfaction vis-à-vis de chaque lieu. Puis, la moyenne des évaluations est calculée pour chaque utilisateur.

Les résultats obtenus sont présentés dans le tableau suivant :

ID utilisateur	moyenne des évaluations	précision
1	3.5000	0.75
3	3.5000	0.5
5	2.0000	0.25
9	3.3333	0.67
10	4.6667	1
11	5.0000	1
12	4.0000	1
13	5.0000	1
14	5.0000	1
15	5.0000	1
16	5.0000	1
17	5.0000	1
18	5.0000	1
19	5.0000	1
20	5.0000	1
21	5.0000	1
22	4.0000	1
23	4.5000	1
24	4.0000	1
25	5.0000	1
26	4.3333	1

TABLE 3.1 – La moyenne des évaluations et des précisions des utilisateurs de système "Findtome".

Nous avons utilisé la mesure de Precision qui s'agit de l'une des mesures les plus utilisées dans l'évaluation des performance des systemes de recommandation. La precision est donnée par la formule (09) ci dessous :

$$P = \frac{\text{nombre de lieux recommandés jugés intéressants}}{\text{le nombre total des lieux recommandés}} [9]$$

Comme le montre le tableau (01), parmi les 21 utilisateurs qui ont évalué les recommandations, nous trouvons que la moyenne des évaluations pour 20 utilisateurs est supérieure à 3.

De plus, la précision dans la majorité des cas est égale à 1. Alors, la précision moyenne calculée est égale à 0.91. Une telle valeur de précision est très encourageante.

Précision moyenne = 0.91

Exemple :

pour l'utilisateur 26 il a donné 3 évaluations positives (nombre d'étoiles supérieur ou égal à 3) aux lieux qu'on lui a recommandé. Alors, la précision donnée par l'équation (09) est comme suit : $3/3=1$.

ID utilisateur	ID place	Évaluation
26	174	4
26	101	5
26	193	4
Total	places	$3/3 = 1$
Précision = 1		

TABLE 3.2 – Calcul de précision pour l'utilisateur numéro 26.

La précision calculée pour les lieux recommandés était 0.91, elle est considérée comme un résultat très encourageant. En effet la majorité des utilisateurs ont été satisfaits des recommandations suggérées par le système. On peut conclure que le système réussit à recommander des lieux intéressants aux utilisateurs, en se basant sur la localisation, la distance, les évaluations (notations et commentaires) et l'âge.

3.5 Conclusion

Dans ce chapitre, nous avons présenté les outils de développement, les étapes d'implémentation et les différentes interfaces de notre système. Nous avons également présenté les résultats des évaluations exprimés en matière de précision.

Le but principal de cette implémentation est d'évaluer notre proposition de système de recommandation de lieux et de valider et justifier les différentes phases proposées afin d'améliorer la qualité de recommandation.

CONCLUSION GÉNÉRALE

Les Systèmes de Recommandation (SR) sont des outils de recherche d'information et de filtrage qui vise à proposer aux utilisateurs des items qui pourraient les intéresser. La plupart des SR se basent sur l'analyse d'historique d'évaluation des items afin de prédire l'intérêt qu'un utilisateur peut porter à un item donné.

Dans ce mémoire, nous nous sommes intéressés au domaine de la recommandation des lieux adéquat aux préférences des différents utilisateurs. En effet, dans ce projet un système de recommandation des lieux intéressants et spécifiques pour chaque utilisateur est développé. Des paramètres contextuels ont été utilisés à cette fin qui sont : la localisation, la distance, les évaluations (notations et commentaires) et l'âge.

Les résultats que nous avons obtenus (en termes de précision) sont encourageants et nous permettent de conclure que notre système de recommandation des lieux réussit à recommander aux utilisateurs des items qui correspondent à leurs préférences.

A long terme, nous envisageons de proposer tout d'abord d'enrichir la collection de données par de nouveaux lieux et leurs évaluations. Aussi on propose de considérer d'autres paramètres contextuels comme le jour (jour de travail ou fin de semaine).

Finalement, ce travail nous a permis d'apprendre beaucoup de choses et d'explorer de nouvelles méthodes dans ce domaine de recherche et de lire plusieurs articles intéressants.

BIBLIOGRAPHIE

- [1] Zahra ABBASI-MOUD, Hamed VAHDAT-NEJAD et Javad SADRI. « Tourism recommendation system based on semantic clustering and sentiment analysis ». In : *Expert Systems with Applications* 167 (2021), p. 114324.
- [2] Larbes ABDELKRIM, Amrani OKBA et Khantoul BILEL. « Une Approche deep learning pour l'analyse des sentiments ». In : (2021).
- [3] Sibel ADALI, Fred SISENDA et Malik MAGDON-ISMAIL. « Actions speak as loud as words : Predicting relationships from social behavior data ». In : *Proceedings of the 21st international conference on World Wide Web*. 2012, p. 689-698.
- [4] Gediminas ADOMAVICIUS et Alexander TUZHILIN. « Toward the next generation of recommender systems : A survey of the state-of-the-art and possible extensions ». In : *IEEE transactions on knowledge and data engineering* 17.6 (2005), p. 734-749.
- [5] Mehdi Hosseinzadeh AGHDAM. « Context-aware recommender systems using hierarchical hidden Markov model ». In : *Physica A : Statistical Mechanics and its Applications* 518 (2019), p. 89-98.
- [6] *ajax*. URL : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203617-ajax-asynchronous-javascript-and-xml-definition-traduction/>.

- [7] Ricardo ANACLETO et al. « Mobile application to provide personalized sight-seeing tours ». In : *Journal of Network and Computer Applications* 41 (2014), p. 56-64.
- [8] Taushif ANWAR et V UMA. « Comparative study of recommender system approaches and movie recommendation using collaborative filtering ». In : *International Journal of System Assurance Engineering and Management* 12.3 (2021), p. 426-436.
- [9] *apify*. URL : <https://www.capterra.fr/software/150854/apify>.
- [10] *appsgeyser*. URL : <https://appsgeyser.com/blog/appsgeyser-android-app-tool-platform/>.
- [11] *bootstrap*. URL : <https://www.lemagit.fr/definition/Bootstrap>.
- [12] Robin BURKE. « Hybrid recommender systems : Survey and experiments ». In : *User modeling and user-adapted interaction* 12.4 (2002), p. 331-370.
- [13] Jean-Jacques CHAROLLAIS et al. « La Molasse du bassin franco-genevois et son substratum ». In : *Archives des Sciences* 60 (2007), p. 59-174.
- [14] Ricardo COLOMO-PALACIOS et al. « Towards a social and context-aware mobile recommendation system for tourism ». In : *Pervasive and Mobile Computing* 38 (2017), p. 505-515.
- [15] CAMELIA DADOUCHI et BRUNO AGARD. « CIGI 2017 État de l'art sur les systèmes de recommandation ». In : ().
- [16] *diagramme de cas d'utilisation*. URL : <https://www.ibm.com/docs/fr/rational-soft-arch/9.5?topic=diagrams-use-case>.
- [17] *diagramme de sequence*. URL : <https://www.ibm.com/docs/fr/rsm/7.5.0?topic=uml-sequence-diagrams>.
- [18] Alexander FELFERNIG et Robin BURKE. « Constraint-based recommender systems : technologies and research issues ». In : *Proceedings of the 10th international conference on Electronic commerce*. 2008, p. 1-10.

- [19] Charles J FRAZIER et al. « Acetylcholine activates an α -bungarotoxin-sensitive nicotinic current in rat hippocampal interneurons, but not pyramidal cells ». In : *Journal of Neuroscience* 18.4 (1998), p. 1187-1195.
- [20] Gina GEORGE et Anisha M LAL. « Review of ontology-based recommender systems in e-learning ». In : *Computers & Education* 142 (2019), p. 103642.
- [21] Khairil Imran GHAUTH et Nor Aniza ABDULLAH. « Learning materials recommendation using good learners' ratings and content-based filtering ». In : *Educational technology research and development* 58.6 (2010), p. 711-727.
- [22] *google colab*. URL : <https://research.google.com/colaboratory/faq.html>.
- [23] *html css*. URL : <https://www.futura-sciences.com/tech/definitions/internet-css-4050/>.
- [24] Clayton HUTTO et Eric GILBERT. « Vader : A parsimonious rule-based model for sentiment analysis of social media text ». In : *Proceedings of the international AAAI conference on web and social media*. T. 8. 1. 2014, p. 216-225.
- [25] IBM. *diagramme de classe*. URL : <https://www.ibm.com/docs/fr/rsar/9.5?topic=diagrams-class>.
- [26] *INFOpedia - VADER Sentiment Analysis*. 2014. URL : <https://cmci.colorado.edu/infopedia/pmwiki.php?n=Analytics.VADERSentimentAnalysis>.
- [27] *javascript*. URL : https://developer.mozilla.org/fr/docs/Learn/JavaScript/First_steps/What_is_JavaScript.
- [28] *JQuery*. URL : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203587-jquery-definition/>.
- [29] T Jaya LAKSHMI et S Durga BHAVANI. « Link Prediction Approach to Recommender Systems ». In : *arXiv preprint arXiv :2102.09185* (2021).
- [30] Marilyne LATOUR. « Analyse de sentiments dans les textes économiques : un exemple d'application chez ReportLinker ». In : ().
- [31] *leaflet*. URL : <https://leafletjs.com/>.

- [32] Stanley LOH et al. « A tourism recommender system based on collaboration and text analysis ». In : *Information Technology & Tourism* 6.3 (2003), p. 157-165.
- [33] Tinghuai MA et al. « Social network and tag sources based augmenting collaborative recommender system ». In : *IEICE transactions on Information and Systems* 98.4 (2015), p. 902-910.
- [34] MySQL. URL : <https://kinsta.com/fr/base-de-connaissances/qu-est-ce-que-mysql/>.
- [35] *numpy*. URL : <https://numpy.org/doc/stable/user/whatisnumpy.html>.
- [36] *pandas*. URL : <https://datascience.eu/fr/programmation-informatique/python-pandas-dataframe/>.
- [37] Michael J PAZZANI et Daniel BILLSUS. « Content-based recommendation systems ». In : *The adaptive web*. Springer, 2007, p. 325-341.
- [38] *php*. URL : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203597-php-hypertext-preprocessor-definition/>.
- [39] Vidisha M PRADHAN, Jay VALA et Prem BALANI. « A survey on sentiment analysis algorithms for opinion mining ». In : *International Journal of Computer Applications* 133.9 (2016), p. 7-11.
- [40] *python*. URL : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1445304-python-definition-et-utilisation-de-ce-langage-informatique/>.
- [41] Paul RESNICK et al. « Grouplens : An open architecture for collaborative filtering of netnews ». In : *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. 1994, p. 175-186.
- [42] Pierre RICCI, Sibylle BUI et Claire LAMINE. *Repenser la protection des cultures : innovations et transitions*. Editions Quae, 2011.
- [43] Gerard SALTON. « Automatic text processing : The transformation, analysis, and retrieval of ». In : *Reading : Addison-Wesley* 169 (1989).

- [44] Claire B SMALLWOOD, Lynnath E BECKLEY et Susan A MOORE. « An analysis of visitor movement patterns using travel networks in a large marine park, north-western Australia ». In : *Tourism Management* 33.3 (2012), p. 517-528.
- [45] *sublime*. URL : <https://kinsta.com/fr/blog/comment-utiliser-sublime-text/>.
- [46] John K TARUS, Zhendong NIU et Abdallah YOUSIF. « A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining ». In : *Future Generation Computer Systems* 72 (2017), p. 37-48.
- [47] Adithya THADURI, Uday KUMAR et Ajit Kumar VERMA. « Computational intelligence framework for context-aware decision making ». In : *International Journal of System Assurance Engineering and Management* 8.4 (2017), p. 2146-2157.
- [48] Gytis TUMAS et Francesco RICCI. « Personalized mobile city transport advisory system ». In : *Information and communication technologies in tourism 2009* (2009), p. 173-183.
- [49] Karsten TYMANN et al. « GerVADER-A German Adaptation of the VADER Sentiment Analysis Tool for Social Media Texts. » In : *LWDA*. 2019, p. 178-189.
- [50] *vader*. URL : <https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>.
- [51] Veena Prasad VEMURI et al. « Analyzing the use of internet of things (IoT) in artificial intelligence and its impact on business environment ». In : *Materials Today : Proceedings* 51 (2022), p. 2194-2197.
- [52] Guocan YU, Kecheng JIE et Feihe HUANG. « Supramolecular amphiphiles based on host-guest molecular recognition motifs ». In : *Chemical reviews* 115.15 (2015), p. 7240-7303.
- [53] Shufa ZHENG et al. « Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020 : retrospective cohort study ». In : *bmj* 369 (2020).

-
- [54] Xiaolin ZHENG et al. « Topic tensor factorization for recommender system ». In : *Information Sciences* 372 (2016), p. 276-293.
- [55] Xiaoyao ZHENG et al. « A tourism destination recommender system using users' sentiment and temporal dynamics ». In : *Journal of Intelligent Information Systems* 51.3 (2018), p. 557-578.