

UNIVERSITE DE BLIDA 1
Faculté des sciences
Département d'informatique



Mémoire de Master
Spécialité : Informatique

Option : Ingénierie des Logiciels

Étude de l'hypothèse de Food Pairing dans la cuisine maghrébine

Réalisé par
ABDALLAH Adel
TALI Fella

Supervisé par
Dr. BACHA Siham

Devant les membres de jury
Président : Blida 1
Examineur : Blida 1

Juin 2023

Résumé

La cuisine est au cœur de notre identité culturelle, et les coutumes culinaires ont toujours été très différentes d'une région à l'autre. La gastronomie computationnelle, un domaine interdisciplinaire qui utilise la science des données pour étudier les aliments, afin d'étudier les relations qui sous-tendent les structures de ces différentes cuisines.

L'hypothèse de food pairing, selon laquelle les ingrédients combinés avec des composés aromatiques similaires ont un meilleur goût, est l'une des questions clés dans ce domaine. Elle a fait l'objet de recherches sur les cuisines occidentales, européennes et moyen-orientales, mais aucune étude n'a été réalisée dans la cuisine maghrébine.

Dans cette étude, nous avons utilisé l'architecture LSTM, pour tester cette hypothèse dans la cuisine maghrébine en utilisant des recettes maghrébines traditionnelles tirées du livre de Mme Salima HADJIAT "la cuisine maghrébine", puis prétraitées à l'aide de diverses méthodes NLP. Notre étude a révélé que les composants utilisés dans les plats maghrébins ont tendance à s'associer négativement. Ce résultat présente une similitude avec la cuisine sud européenne et indique une tendance régionale méditerranéenne.

Mots clés : Food Pairing, cuisine maghrébine, traitement automatique des langues

المخلص

المطبخ هو في الصميم هويتنا الثقافية ، و كانت عادات الطهي دائماً مختلفة تماماً من منطقة الى اخرى. فن الطهو الحسائي، و هو مجال متعدد التخصصات يستخدم علم البيانات لدراسة العلاقات التي تكمن وراء هياكل هذه الماكولات المختلفة. فرضية إقران الطعام تقول ان الاكل المكون من مواد و عناصر عطرية متشابهة له ذوق افضل. تم تطبيق هذه الفرضية على الماكولات الغربية و الاوروبية و الشرق الاوسط، و لكن لم يسبق تجريبه على الاكل المغاربي. في هذه الدراسة، استخدمنا بنية الذاكرة طويلة قصيرة المدى، و هي خوارزمية جينية، لاختبار هذه الفرضية في المطبخ المغاربي باستخدام وصفات مغربية تقليدية من كتاب السيدة سليمة محيات المطبخ المغاربي، ثم تمت معالجتها مسبقا باستخدام طرق البرمجة اللغوية العصبية المختلفة. وجدت دراستنا ان المكونات المستخدمة في الاطباق المغاربية تميل الى الارتباط سلبا بالاطعمة الاخرى. تتوافق هدم النتيجة مع المطبخ الجنوب الاوروي هذا يوضح طبع منطقة البحر الابيض المتوسط.

الكلمات المفتاحية : إقران الطعام ، المطبخ المغاربي ، معالجة اللغة الطبيعية

Remerciements

Ce présent travail de mémoire qui marque la fin de nos études en ingénierie des logiciels n'a pu aboutir que grâce à la conjugaison des efforts de plusieurs personnes.

Qu'il soit permis d'exprimer notre profonde gratitude à tous ceux qui nous ont encouragés à poursuivre nos études, ceux qui nous ont accompagnés dans ce parcours que ce soit dans notre travail de recherche jusqu'au terme du présent mémoire.

A notre dieu tout puissant, qui nous a comblés de sa force et son intelligence.

Nos remerciement de gratitude s'adressent à notre promotrice Dr.Siham Bacha pour avoir accepté de s'investir corps et âme avec rigueur scientifique pour la direction de ce projet de fin d'étude.

Nous tenons à remercier toutes nos familles et nos amis qui nous ont aidés tant moralement que matériellement.

Dédicace

Je dédie ce travail

Mes parents, qui ont toujours été ma source d'inspiration et de soutien inconditionnel. Vos encouragements et vos sacrifices ont été les fondations solides sur lesquelles j'ai construit mon parcours académique. Je vous suis infiniment reconnaissant.

A mes frères Salim et Mounir, ainsi que ma belle-sœur Sabeha et ma petite nièce meriouma, Je vous suis reconnaissant pour votre présence à chaque étape de ma vie.

Une mention spécial a mes amis qui ont contribué à mon travail : Khaled, Ayoub, Moumen et Alaa vous m'avez fait gagnée énormément de temps j'en suis très reconnaissant.

A mes amies fidèle qui ont toujours été présents a mes côtés, je souhaite citer Yacine mingueza, Hocine moumnin, Chiraz selena, Mohamed lewandowski et Mohamed loutchy et tous mes amis du club ITC et à tous mes amis que je n'ai pas mentionnés mais dont je suis extrêmement reconnaissant et bien sur une spécial dédicace a xvxxii. Votre amitié et votre soutien inconditionnel ont apporté de la joie et du réconfort à ma vie.

Et à ma binôme Fella je tiens à exprimer ma profonde gratitude pour ton soutien moral, ta patience tout au long de notre collaboration.

Adel

Table des matières

Liste des tableaux	ix
Table des figures	x
Introduction générale	1
1 Food Pairing	4
1.1 Contexte	4
1.2 gastronomie computationnelle	5
1.2.1 Food pairing	5
1.2.2 Flavor data	6
1.3 Problématique	7
1.4 Travaux existants	8
1.4.1 Travaux existants sur le food pairing	8
1.4.2 Travaux existants dans le processus de génération	12
1.5 Conclusion	17
2 Approche proposée	18
2.1 Introduction	18
2.2 Schéma global	18
2.3 La collecte des données	20
2.4 Prétraitement	20
2.4.1 Nettoyage des données	20
2.4.2 l'intégration des données	20
2.4.3 Transformation des données	20
2.5 Génération de recettes aléatoires	21

2.5.1	Le processus de génération	21
2.6	Calcul du partage moyen des saveurs	26
2.7	Tester l’hypothèse de food pairing	27
2.8	Conclusion	27
3	Test et validation	28
3.1	Introduction	28
3.2	Environnement de travail :	28
3.2.1	Python	28
3.2.2	Bibliothèque d’expressions régulières	28
3.2.3	Bibliothèque des Pandas	29
3.2.4	Bibliothèque NLTK	29
3.2.5	Bibliothèque Deep translator	29
3.3	Bibliothèque Wordcloud	29
3.4	Bibliothèque Plotly	29
3.5	Bibliothèque NumPy	30
3.6	Ensemble de données :	30
3.6.1	Recettes traditionnelles maghrébines	30
3.6.2	Ensemble de données moléculaires :	31
3.7	Pré traitement	31
3.7.1	Nettoyage des recettes :	31
3.7.2	Lier les ingrédients recettes à leurs profils moléculaires	34
3.8	Visualisation	36
3.9	Processus de génération	37
3.9.1	Génération de recettes synthétiques	37
3.9.2	Résultats de l’algorithme génétique	38
3.9.3	Tests et résultats de la génération aléatoire	40
3.9.4	Tests et résultats de l’architecture LSTM	42
3.10	Tests et résultats de l’architecture LSTM : étude de l’impact des épices . .	45
3.11	L’étude du food pairing dans la cuisine algérienne	48
3.11.1	Résultats de l’hypothèse du food pairing sur ensemble 1 de donnée algérienne	48
3.11.2	Résultats de l’hypothèse du food pairing sur ensemble 2 de donnée algérienne	50
3.12	Conclusion	52

Conclusion générale	53
Références	55

Liste des tableaux

2.1	Annotations des formules ci-dessus et leur description	26
3.1	Liste des recettes remplacées par la liste de leurs ingrédients	33
3.2	Liste des épices composées remplacées par leurs compositions	33
3.3	Liste des recettes où la matière grasse est facultative	36
3.4	Liste de recette où l'ingrédient variantes est facultative	36
3.5	Liste de recette où l'ingrédient absinthe est facultative	36
3.6	Les résultats de GA avec différentes configurations sur les données de la cuisine maghrébine	38
3.7	Les résultats de génération aléatoire avec différentes configurations sur les données de la cuisine maghrébine	40
3.8	Les résultats de LSTM, taille recette 4, 11 et 41, cuisine maghrébine	42
3.9	Les résultats de LSTM, taille de séquence N=4, cuisine maghrébine	44
3.10	Le résultat avec la composition d'ersa fixée sur la cuisine algérienne	45
3.11	Le résultat de la fixation des ingrédients de la chermoula sur la cuisine algérienne	46
3.12	Le résultat de l'impact de la sauce rouge sur la cuisine algérienne	47
3.13	Le résultat de food pairing avec la sauce blanche algérienne	47
3.14	Les résultats des algorithmes génétiques sur l'ensemble de données algériennes extraites du livre de Mme Hadjiat	49
3.15	Les résultats obtenus par la génération aléatoire sur l'ensemble de recettes algériennes de Mme Hadjiat	49
3.16	Les résultats de l'algorithme LSTM sur l'ensemble de données algériennes extraites du livre de Mme Hadjiat	50

Table des figures

1.1	Cartographie	5
1.2	Flavordb details [1]	7
1.3	Réseau de saveurs bipartite [2]	9
1.4	L'épine dorsale du réseau de saveurs [2]	10
1.5	L'organisation des données recueillies dans l'étude la cuisine saudienne[3] .	12
2.1	Aperçu de l'approche proposée	19
2.2	Diagramme illustrant le prétraitement des données	21
2.3	Figure illustrant l'architecture du lstm	22
2.4	Porte d'oubli	23
2.5	Porte d'entrée	24
2.6	Porte de sortie	25
3.1	Diagramme de nombre de recettes et d'ingrédients uniques pour la cuisine maghrébine et algérienne	30
3.2	Diagramme de nombre de recettes et d'ingrédients uniques pour la cuisine tunisienne et marocaine	31
3.3	L'ensemble des données est affiché sous la forme d'un nuage de mots avant le prétraitement.	34
3.4	Nuage de mots affichant la base de données après le prétraitement.	34
3.5	La fréquence des ingrédients dans la cuisine maghrébine	37
3.6	Résultats de GA avec population initiale = 3, cuisine maghrébine	39
3.7	Résultats de GA avec population initiale = 5, cuisine maghrébine	39
3.8	Résultats de GA avec population initiale = 7, cuisine maghrébine	39
3.9	Résultats de GA avec une population initiale = 9, cuisine maghrébine . . .	40
3.10	Résultats de génération aléatoire avec une taille de recette = 4, cuisine maghrébine	41

3.11 Résultats de génération aléatoire avec une taille de recette = 11, cuisine maghrébine	41
3.12 Résultats de génération aléatoire avec une taille de recette = 41, cuisine maghrébine	41
3.13 Résultats de LSTM pour une taille de recette = 4, cuisine maghrébine . . .	43
3.14 Résultats de LSTM pour une taille de recette = 11, cuisine maghrébine . .	43
3.15 Résultats de LSTM pour une taille de recette = 41, cuisine maghrébine . .	43
3.16 Les résultats de LSTM, taille de séquence N=4, taille de recette = 11 et 41 cuisine maghrébine	44
3.17 Résultat pour une taille de recette = 11, composition d'ersa fixée dans les recettes générées	45
3.18 Résultat pour l'utilisation de la chermoula avec une taille de recette = 11 .	46
3.19 Résultat pour une taille de recette = 11, composition de la sauce rouge fixée	47
3.20 Résultat de l'utilisation de la sauce blanche avec une taille de recette = 11	48
3.21 Résultats des GA sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une population initiale = 3 et 5	50
3.22 Résultats des GA sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une population initiale = 7 et 9	51
3.23 Résultat de la génération aléatoire sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une taille de recette = 11	51
3.24 Résultat de l'algorithme LSTM sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une taille de recette = 11	51

Introduction générale

Les aliments que nous mangeons peuvent avoir un impact direct sur notre capacité à profiter pleinement de la vie. Le plaisir représente une motivation principale à l'attrait à nourriture. La consommation et la préparation sont influencées par une variété de facteurs, y compris des considérations culturelles, historiques, religieuses, géographiques et climatiques. La cuisine a été un processus d'apprentissage par essais et erreurs pendant des siècles, mais une fois que les matériaux disponibles ont été maîtrisés, c'est devenu un art qui a ouvert la porte à la créativité. Ces dernières années, les scientifiques ont collaboré avec des gastronomes pour ajouter une couche rationnelle à la compréhension des raisons derrière certaines combinaisons étranges mais savoureuses, ou ce qui génère des goûts et des affinités diverses parmi les populations du monde. La plus récente est l'hypothèse du food pairing de Heston Blumenthal, qu'il a mise en œuvre en tant que chef du restaurant The Fat Duck et pionnier de la cuisine multisensorielle, de food pairing et de l'encapsulation des saveurs. L'idée de base derrière le food pairing est que les éléments qui ont des structures moléculaires similaires interagissent bien. Les scientifiques des données ont utilisé la technologie pour explorer cette notion et proposer des solutions qui aideront les cuisiniers dans leur cheminement créatif.

L'approche culinaire connue sous le nom de "food pairing" gagne en popularité car elle permet de comprendre les combinaisons d'ingrédients sur la base de leurs composants moléculaires et aromatiques. La majorité des études sur la relation entre les aliments se concentrent sur les cuisines européennes et asiatiques. Cependant, malgré l'importance de la nourriture maghrébine dans la gastronomie mondiale, aucune recherche n'a été effectuée sur l'application de l'hypothèse du "Food Pairing" à cette cuisine. Les ingrédients utilisés dans la cuisine maghrébine ont des saveurs et des arômes distinctifs qui proviennent de leur histoire et de leur environnement. Il est donc essentiel de comprendre comment ces ingrédients peuvent être utilisés. Prenons l'exemple d'un restaurant moléculaire réputé appelé "elBulli"¹ situé en Espagne. Ce restaurant s'est efforcé d'offrir une expérience culinaire exceptionnelle en repoussant les limites de la gastronomie traditionnelle. ElBulli a mis en pratique le concept du "food pairing" en utilisant des techniques de cuisine moléculaire pour créer des plats innovants et surprenants. Ils ont expérimenté avec des combinaisons inattendues d'ingrédients, utilisé des techniques de sphérification, de gélification et de mousse pour transformer la texture et la présentation des aliments, offrant ainsi une expérience gustative et visuelle unique. Une étude l'année dernière a été menée sur la cuisine algérienne, qui a jeté les bases pour une compréhension approfondie des combinaisons

1. restaurant situé à Roses, en Catalogne, en Espagne.

d'ingrédients et des saveurs dans cette cuisine spécifique. Cependant, il est important de noter que la cuisine maghrébine englobe un ensemble plus large de traditions culinaires, qui inclut également les cuisines marocaine et tunisienne. En appliquant l'hypothèse du "Food Pairing" à la cuisine maghrébine dans son ensemble, nous pourrions non seulement approfondir nos connaissances culinaires, mais aussi promouvoir et préserver cette riche tradition culinaire pour les générations futures.

L'industrie de la restauration joue un rôle économique crucial, générant des emplois et contribuant à la croissance économique. Elle représente le plus grand segment d'emploi dans le commerce de détail au monde. Aux États-Unis, les ventes d'aliments et de boissons dépassent 700 milliards de dollars, et lorsqu'elles sont combinées avec les supermarchés de détail, elles atteignent plus de 1,2 billion de dollars [4]. Cela met en évidence l'importance économique de l'industrie alimentaire et l'opportunité pour les acteurs de l'industrie de rechercher de nouvelles idées et de nouvelles saveurs pour attirer les consommateurs et rester compétitifs sur le marché mondial.

La gastronomie moléculaire, définie par Hervé This, englobe les aspects sociaux, artistiques et techniques de l'activité culinaire. Ses objectifs sont de comprendre les proverbes culinaires, d'explorer les recettes existantes, d'introduire de nouveaux outils et méthodes culinaires, d'inventer de nouveaux plats, et d'utiliser la gastronomie moléculaire pour sensibiliser le grand public à l'apport de la science dans notre société. Les domaines d'investigation de cette discipline incluent l'étude des transformations des ingrédients par différentes méthodes de cuisson, la perception multisensorielle de la nourriture, les mécanismes de libération des arômes, les préférences gustatives, l'impact des techniques de cuisson sur la texture et la saveur des aliments, l'interprétation des signaux sensoriels par le cerveau, et les influences environnementales et émotionnelles sur notre plaisir de manger. En nous basant sur ces principes

En nous basant sur ces principes, notre mémoire se concentre sur l'étude de recherche pour déterminer le "food pairing" de la cuisine maghrébine. Pour ce faire, nous avons compilé des recettes traditionnelles et authentiques, et nous avons construit le profil gustatif de chaque ingrédient présent dans le corpus, en identifiant leurs molécules aromatiques. Notre premier objectif est d'extraire les ingrédients les plus couramment utilisés dans la cuisine maghrébine en utilisant des techniques de traitement automatique du langage naturel et des outils de visualisation. Cependant, nous avons rencontré plusieurs obstacles au cours de notre recherche, notamment le manque de documentation sur la culture culinaire maghrébine et l'absence d'un corpus de recettes authentiques représentatives. Cela nous a obligés à créer manuellement notre propre corpus de recettes à travers un livre, ce qui a été coûteux en termes de temps, car nous n'avons pas trouvé de version PDF. Le manque de composants moléculaires dans certains ingrédients a également posé un autre défi dans notre processus. Pour résoudre ce problème, nous avons utilisé un algorithme LSTM pour générer des recettes et ensuite étudier les résultats du "food pairing".

L'objectif de notre mémoire est de faire une étude sur le food pairing sur la cuisine maghrébine en utilisant des méthodes basées sur le traitement automatique du langage naturel, l'analyse moléculaire, LSTM.

On va entamer notre premier chapitre en présentant les concepts clés qui sous-tendent

notre recherche, tels que food pairing, flavor data, ainsi que notre état de l'art sur les travaux précédents dans le domaine. le chapitre suivant est dédié à la conception de notre solution, tandis que le troisième chapitre traite de l'implémentation, les tests et la validation de notre approche.

Chapitre 1

Food Pairing

1.1 Contexte

La préparation d'un repas est comme raconter une histoire. Chaque ingrédient et chaque technique de cuisine ont une raison d'être et une logique. C'est une pratique courante dans de nombreuses cultures, car la nourriture est un langage universel qui permet aux gens de communiquer et de s'exprimer. Depuis toujours, la nourriture est un élément fondamental à évaluer en fonction de facteurs environnementaux, sociaux et culturels. La cuisine d'un pays reflète son histoire, ses coutumes, sa géographie et son climat, ainsi que les croyances et les valeurs de son peuple. Au fil des ans, la cuisine est devenue un art qui permet des créations culinaires innovantes et extraordinaires. Pour rationaliser la nourriture, les scientifiques ont collaboré avec des chefs, utilisant la chimie et la technologie informatique pour comprendre les relations et les préférences gustatives des gens.

Dans ce contexte, la gastronomie moléculaire a émergé comme une discipline scientifique passionnante qui explore les principes physico-chimiques sous-jacents à la préparation des aliments, ouvrant ainsi la voie à des expérimentations culinaires innovantes et à la création de nouvelles sensations gustatives. La théorie du "food pairing" est basée sur l'idée que les composants aromatiques de deux aliments qui ont quelque chose en commun les rendent plus savoureux ensemble. Des chercheurs du monde entier ont cherché à confirmer ou à infirmer cette hypothèse, et il s'avère que certaines combinaisons alimentaires improbables mais délicieuses sont effectivement liées à des composés aromatiques communs.

Le food pairing est une théorie qui a été étudiée pour de nombreuses cuisines à travers le monde, notamment en Europe, en Asie et même en Arabie Saoudite. L'année dernière, une analyse avait été menée sur la gastronomie Algérienne, les résultats étaient négatifs. Notre but est de poursuivre la cartographie de l'hypothèse du food pairing mondiale en étudiant la cuisine du Maghreb comme le montre la figure 1.1, en incluant cette fois les cuisines du Maroc et de la Tunisie. Nous espérons acquérir une compréhension plus complète de cette pratique culinaire. En incluant ces deux pays dans notre étude, nous pouvons également voir si les tendances sont similaires ou différentes de l'Algérie. Cela nous aidera à mieux comprendre la relation entre les ingrédients et à identifier les combinaisons les plus délicieuses de la cuisine maghrébine.

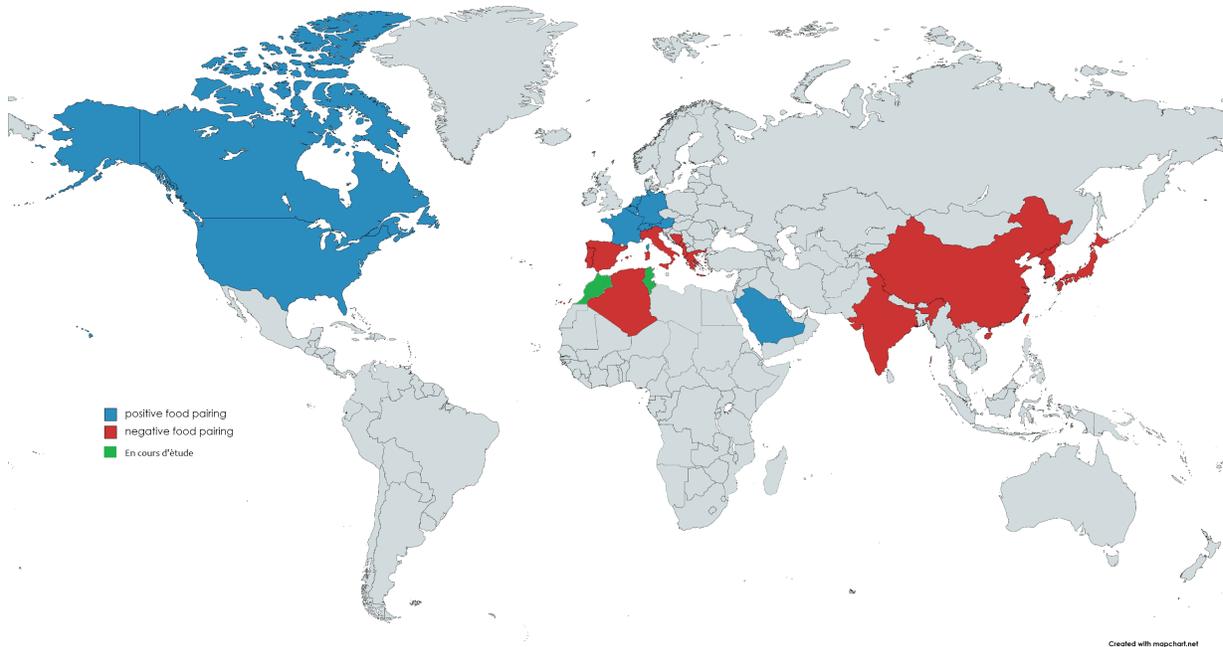


FIGURE 1.1 – Cartographie

1.2 gastronomie computationnelle

La gastronomie en tant que terme est l'art ou la science de bien manger[5], en 2013, Ahnert a introduit le terme « gastronomie computationnelle »[6], faisant référence à un nouveau domaine apparu en raison du récent déluge de données sur la composition des aliments, la composition chimique et les combinaisons d'appariements possibles. La fonction des techniques de fouille de données et d'extraction de connaissances a davantage favorisé le développement de ce domaine. À notre connaissance, il n'existe aucune explication scientifique aux saveurs et arômes uniques et attrayants des différentes cuisines du monde. Une explication possible, cependant, est que les composés volatils dans les aliments stimulent les récepteurs dans la bouche et le nez, créant une expérience alimentaire unique. Ainsi, les produits chimiques dans les aliments semblent jouer un rôle important dans la définition des goûts et des arômes que nous associons aux différentes cuisines[7]. Il existe différentes dimensions de la gastronomie computationnelle comme : l'hypothèse du food pairing et base de donné moléculaire.

1.2.1 Food pairing

La nourriture est un symbole de la vie elle-même, elle est à la fois essentielle et complexe. Depuis l'Antiquité, les humains ont été poussés à explorer et à apprendre de nouvelles choses par leur amour de l'apprentissage. L'une d'entre elles consiste à jouer avec les aliments pour créer de nouvelles saveurs. Une fraise aura une saveur particulière lorsqu'elle est mangée avec du chocolat, mais une fraise mangée avec du fromage aura une toute autre

saveur[7] , par contre ils n’ont pas encore exploré toutes les combinaisons possibles d’aliments dans les différentes cuisines régionales [2]. ”Food pairing” est l’idée qui sous-tend ces combinaisons qui a été découverte en 1992 par Heston Blumenthal, le chef renommé du célèbre restaurant Fat Duck 3. Cette hypothèse stipule que les aliments contenant un grand nombre de composés aromatiques communs ont un goût délicieux lorsqu’ils sont combinés. Cette hypothèse a été étudiée en relation avec leur alimentation par des chercheurs de différents horizons culturels. Certaines cultures, comme la cuisine occidentale, ont obtenu des résultats positifs, ce qui signifie que les ingrédients utilisés dans les recettes de ces cuisines ont tendance à partager davantage de composés aromatiques. L’inverse est vrai pour les cuisines de l’Asie de l’Est et de l’Inde, où les ingrédients ayant plus de composants chimiques communs sont moins susceptibles d’être utilisés dans la même recette[7].

1.2.2 Flavor data

Pour réaliser une étude approfondie sur le food pairing en cuisine maghrébine, il est essentiel de comprendre les bases moléculaires des ingrédients qui composent nos plats. Pour cela nous utilisons une base de données moléculaire ”Flavor DB” [8] qui est un référentiel complet de composée aromatiques, structure physico-chimique, sa source naturelle et son arôme[9]. Il aide à trouver les molécules gustatives désirables, examiner les molécules des ingrédients pour découvrir les aliments apparentés, connaître la chimie des mélanges et des senteurs. La base contient 25 595 molécules aromatiques, 34 catégories d’ingrédients couvrant 936 ingrédients liés à 527 ressources naturelles uniques.

C’est une base de données contenant une large variété de molécules pour chaque ingrédient, il contient la plupart des ingrédients que nous utilisons, Cependant malgré nos recherches et la popularité de la cuisine maghrébine, certains ingrédients ne sont pas présents dans la base de données.

Dans ce cas la nous avons dû remplacer ces ingrédients par d’autres similaires ou la supprimer de la recette si nous avons déterminé qu’ils n’affectent pas significativement les saveurs de la préparation. Nous avons également veillé à ce que les substitutions ne modifient pas le profil moléculaire de la recette de manière significative, afin de préserver son authenticité. (figure1.2)

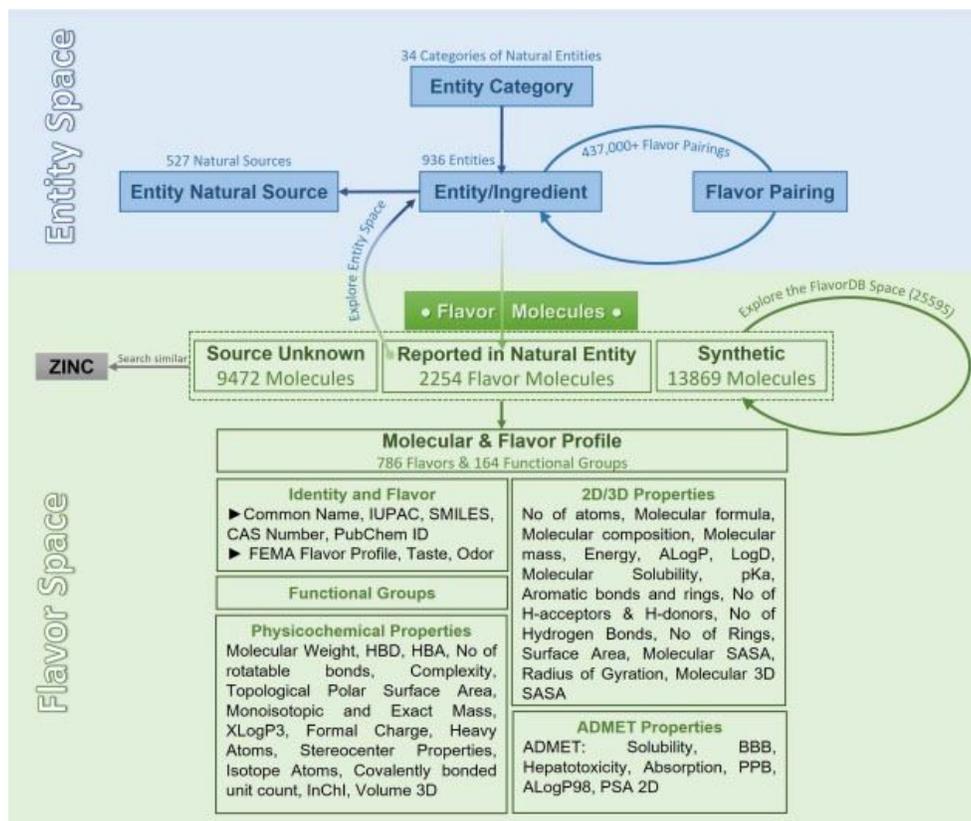


FIGURE 1.2 – Flavordb details [1]

1.3 Problématique

Une approche culinaire appelée "food pairing", qui fonde sa compréhension des combinaisons d'ingrédients sur leurs composants moléculaires et aromatiques, devient de plus en plus populaire. La majorité des recherches sur l'association des aliments se concentre sur les cuisines asiatiques et européennes. Par exemple, Varshney et al. (2013) ont étudié des recettes médiévales européennes et ont constaté que les ingrédients étaient fréquemment combinés en raison de leurs propriétés aromatiques communes[10]. En outre, Jain et Bagler (2015) ont étudié les épices indiennes et démontré que la façon dont elles sont utilisées dans la cuisine indienne repose sur des associations qui renforcent la saveur et l'arôme de chaque plat[11]. En revanche, malgré la popularité de la cuisine maghrébine dans la gastronomie mondiale, aucune étude n'a été réalisée sur l'utilisation de l'hypothèse du "Food Pairing" dans cette cuisine. Les ingrédients utilisés dans la cuisine maghrébine ont des saveurs et des arômes distincts qui proviennent de leur histoire et de leur environnement. Il est donc essentiel de comprendre comment ces ingrédients peuvent être combinés pour créer des plats savoureux et équilibrés. Alors, Qu'en est-il de la nourriture maghrébine? Les habitants du Maghreb préfèrent-ils les saveurs ayant des goûts chimiques communs, ou est-ce l'inverse? Quels sont les facteurs qui ont influencé la culture culinaire maghrébine? L'influence coloniale est-elle significative, ou est-elle plus influencée par sa situation géographique?

1.4 Travaux existants

1.4.1 Travaux existants sur le food pairing

En 1992, le célèbre chef Heston Blumenthal du restaurant The Fat Duck cherchait un substitut au sel dans les recettes de chocolat blanc. Au cours de ces recherches, il fait une découverte surprenante : l'association du caviar et du chocolat blanc produit une délicieuse mixture. Inspiré par ce succès, il décide de rendre visite à François Benzi, scientifique spécialiste des arômes et des parfums, pour découvrir les raisons de cette harmonie gastronomique. Leurs recherches ont montré que le caviar et le chocolat contenaient de grandes quantités d'amines, que l'on trouve dans le fromage, le bœuf séché et d'autres aliments. Le succès de leur mariage semble être en grande partie dû au partage de cette protéine [12]. C'est ainsi qu'est née l'hypothèse du "food pairing" selon laquelle les recettes contiennent souvent des ingrédients de compositions chimiques communes. Depuis, de multiples études ont été menées sur différentes cuisines à travers le monde pour prouver ou infirmer cette hypothèse. Une des études les plus récentes et importantes dans ce domaine est présentée dans [2] et aborde le problème en utilisant une approche topologique. Pour tester cette hypothèse et trouver des systèmes qui vont au-delà de plats spécifiques, les chercheurs ont créé un réseau de saveurs binaires qui met en évidence les composants chimiques partagés par les ingrédients. Ils ont utilisé un vaste ensemble de données de 56 498 recettes provenant de sources de référence américaines, auxquelles ils ont ajouté des références coréennes pour éviter les interprétations occidentales des régimes alimentaires mondiaux. Une recette moyenne contient environ 8 ingrédients. Les recettes sont organisées en cinq catégories géographiques pour fournir un aperçu des cuisines mondiales : Amérique du Nord, Europe de l'Ouest, Europe du Sud, Amérique latine et Asie de l'Est. Ces catégories géographiques sont présentées dans la partie C de la figure 1.3 .

Afin de l'adapter à la recherche, certaines modifications ont été apportées à la base de données. Par exemple, les huiles essentielles ou les extraits ont été remplacés par des composés aromatiques présents dans l'ingrédient d'origine. De plus, les composants de saveur de composants plus spécifiques sont combinés avec des composants plus généraux lors du pré traitement des données. Par exemple, les substances aromatiques présentes dans la viande sont ajoutées au porc et au bœuf. Pour mener à bien cette étude, les chercheurs ont d'abord créé un réseau biparti composé de deux types de nœuds différents, comme le montre la partie A de la figure 1.3 . Le premier nœud contient 381 ingrédients utilisés dans les recettes du monde entier. Le deuxième nœud comprend 1021 composés aromatiques qui constituent les ingrédients utilisés sur le premier nœud. Comme le montre la figure 1.3.A, chaque ingrédient est associé à son composé aromatique respectif. Le poids de chaque lien représente le nombre de composés partagés, comme le montre la figure 1.3.B

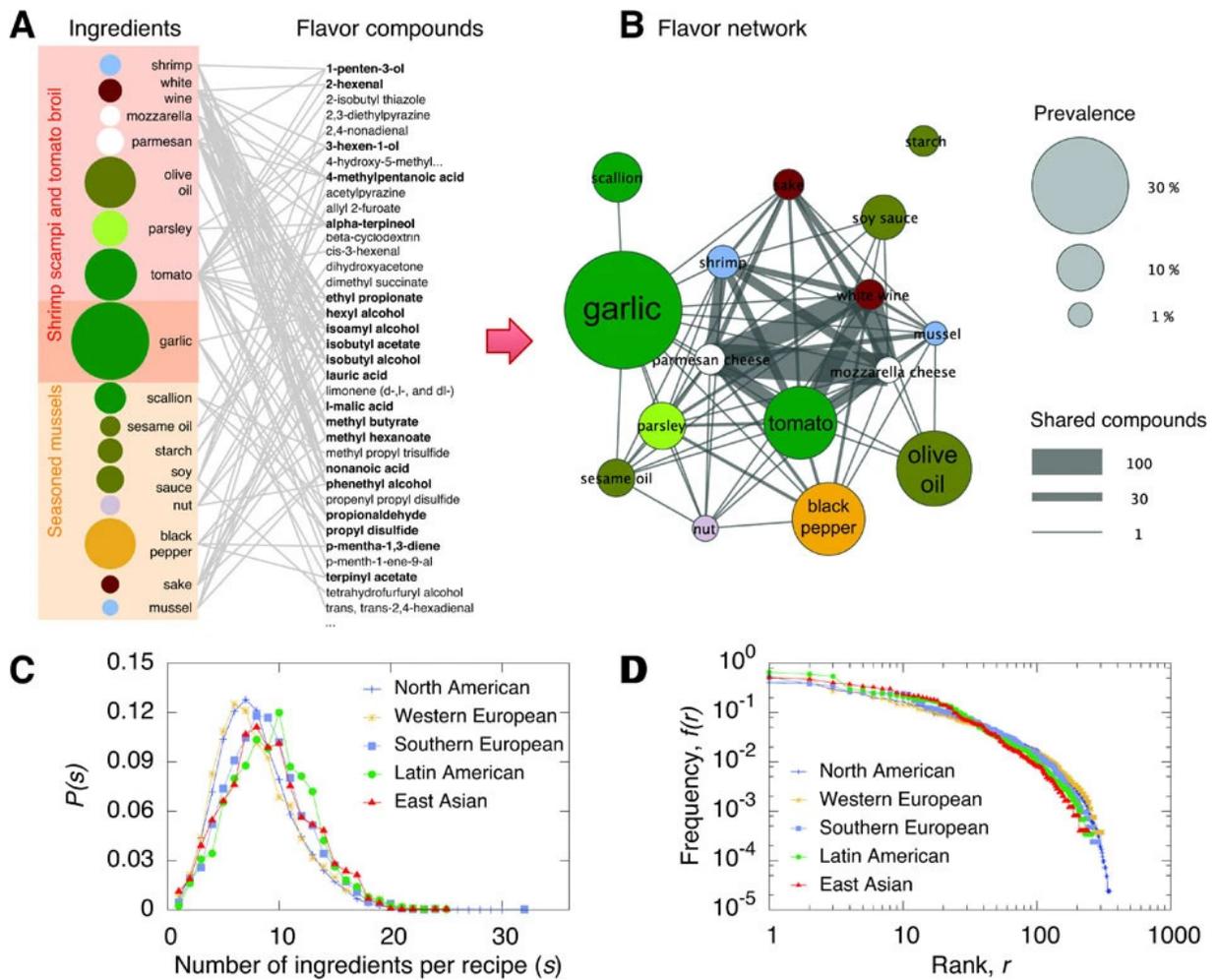


FIGURE 1.3 – Réseau de saveurs bipartite [2]

La figure 1.3.D montre que les cuisines d’Amérique du Nord et d’Europe occidentale préfèrent présenter des recettes utilisant des ingrédients qui partagent des composés aromatiques. À l’inverse, les cuisines d’Asie de l’Est et d’Europe du Sud évitent d’utiliser des ingrédients qui partagent ces composés aromatiques.

Bien que de nombreux ingrédients partagent plusieurs composés aromatiques, le réseau d’arômes est trop petit pour afficher tous les ingrédients. Par conséquent, une méthode d’extraction a été utilisée pour déterminer les liens entre chaque ingrédient. Les relations entre les différents ingrédients sont illustrées dans la figure 1.4, où la taille de la relation indique la quantité de composés aromatiques partagés.

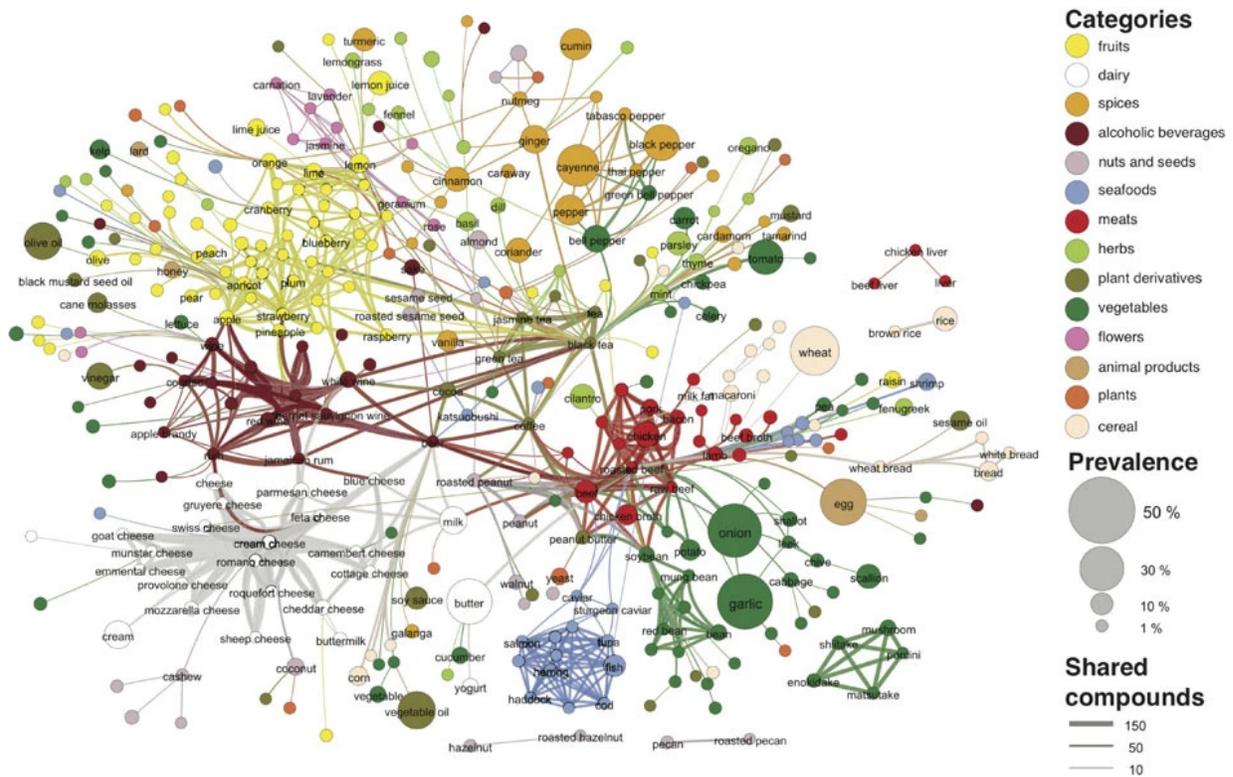


FIGURE 1.4 – L'épine dorsale du réseau de saveurs [2]

L'étude menée par Kush et al[10] se concentre sur la nourriture européenne médiévale et sa propension à mélanger les saveurs. Les chercheurs ont compilé une base de données massive d'informations sur les recettes européennes médiévales couvrant les années 1300 à 1615 en utilisant 25 sources textuelles d'Angleterre, d'Allemagne, de France et d'Italie. Ils ont pris les ingrédients de ces recettes et ont créé manuellement un tableau des ingrédients en tenant compte de la fréquence des mots, des synonymes et des variantes orthographiques. Après avoir prétraité les données, ils ont pu éliminer les recettes vides pour obtenir 4 126 recettes utilisables. L'objectif principal de l'étude était de déterminer comment ces différences dans les combinaisons d'arômes affectaient le produit final. Afin d'établir une corrélation entre les ingrédients des recettes et les composés aromatiques provenant de diverses sources de données, comme Fenaroli et la VCF, les chercheurs ont comparé les données des recettes médiévales avec celles des composés aromatiques. Les conclusions tirées de ces bases de données étaient contradictoires : selon Fenaroli, la cuisine médiévale européenne tendait fortement à mélanger les saveurs, tandis que la source de données VCF montrait une tendance à la séparation des saveurs. Les chercheurs ont étudié cette divergence en examinant les contributions spécifiques de chaque ingrédient et en constatant des variations dans les principaux contributeurs en fonction de la base de données utilisée. Malgré ces contradictions, les auteurs sont parvenus à la conclusion que, dans l'ensemble, la cuisine européenne présentait une tendance positive au mélange des goûts, même si ce mouvement était peut-être plus visible à l'époque en raison de la rareté des ingrédients disponibles. Cette étude a inspiré d'autres recherches dans le domaine des mélanges de saveurs et a révélé d'autres raisons qui expliquent cette tendance.

Une autre étude[11] s'est concentrée sur la cuisine indienne et a présenté un modèle qui quantifie le schéma de combinaison des saveurs et identifie les caractéristiques statistiques de cette cuisine. Les chercheurs ont recueilli 3330 recettes indiennes authentiques dans des livres et des archives en ligne, qu'ils ont ensuite filtrées en éliminant les doublons et les ingrédients sans composés aromatiques, en substituant certains ingrédients répétitifs (ananas in preserve = ananas), en éliminant les ingrédients des catégories "en-cas" et "additifs", et enfin en excluant les recettes ne contenant qu'un seul ingrédient. En suivant ces étapes, ils ont divisé les 2 543 recettes restantes en huit cuisines régionales : Bengali, Gujarati, Jain, Maharashtrian, Mughlai, Punjabi, Rajasthani et Inde du Sud. Ils ont pu obtenir un total de 192 ingrédients, répartis en 15 catégories d'aliments, dont les épices, les légumes, les fruits, le poisson et d'autres produits d'origine animale etc. . .

Des recherches récentes utilisent des algorithmes génétiques pour étudier la cuisine saoudite[3]. L'objectif de cette étude, publiée en 2017 et première du genre dans le monde arabe, était de déterminer si le mélange des saveurs dans la cuisine saoudite avait des effets bénéfiques ou négatifs. Cette étude propose une méthode d'analyse de données et de modélisation pour étudier les aliments à travers leurs composants chimiques. La première étape de leur travail a été la collecte de données, qui a consisté à compiler une centaine de recettes tirées de livres de cuisine traditionnelle saoudite tels que "Les principes de cuisine de Rabha Hafzi pour la cuisine saoudite et moyen-orientale". Pour déterminer le profil d'un ingrédient, une liste d'ingrédients est compilée à partir de chaque recette et, pour chaque ingrédient, une liste de composés aromatiques est compilée à l'aide de la source de données "The Manual of Aromatic Compounds of Fenaroli" (Manuel des composés aromatiques de Fenaroli). La figure 1.5 illustre l'organisation des données recueillies dans le cadre de cette étude.

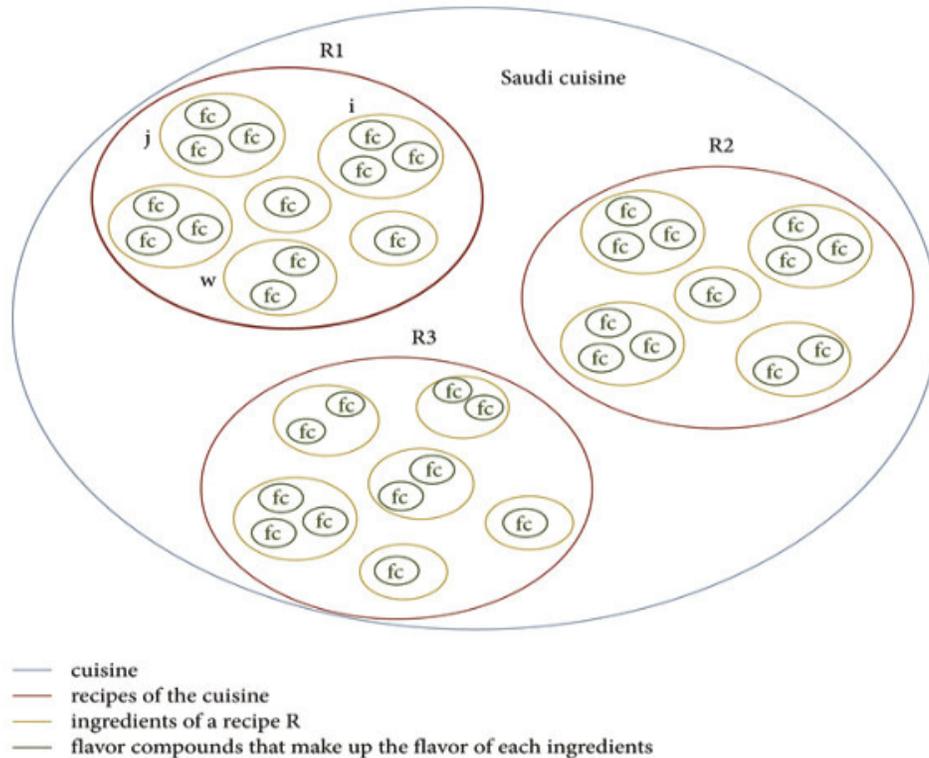


FIGURE 1.5 – L’organisation des données recueillies dans l’étude la cuisine saudienne[3]

1.4.2 Travaux existants dans le processus de génération

La génération automatique de contenu est un domaine de recherche en constante évolution qui a suscité un grand intérêt ces dernières années. Les algorithmes d’apprentissage automatique tels que les réseaux de neurones, les transformers et les GANs peuvent être utilisés pour générer divers types de contenu, tels que des images, des textes, des séquences d’ADN, des protéines, etc. L’objectif de ce travail est de réaliser un état de l’art sur les méthodes de génération de contenu en utilisant différentes approches d’apprentissage automatique. Nous analysons les travaux précédents qui ont utilisé ces différentes méthodes pour générer du contenu dans divers domaines, et évaluons leur potentiel pour la génération de nouveau contenu.

Les GANs (Generative Adversarial Networks) sont des algorithmes de machine learning qui sont utilisés pour générer des données réalistes en simulant un processus de création. Dans le domaine de la cuisine, les GANs sont de plus en plus utilisés pour générer des recettes de cuisine nouvelles et originales.

- Amaia Salvador, Michal Drozdal . . . al [13] ils ont décrit une méthode “Learning to Cook with Synthetic Data” pour la génération de recettes de cuisine à l’aide de données synthétiques. Ils proposent un modèle de réseau de neurones appelé CookGAN, qui utilise une architecture de génération de texte et une méthode d’apprentissage par renforcement pour apprendre à générer des recettes. Le modèle est entraîné sur une grande quantité de

données de recettes existantes pour apprendre les motifs et les règles courantes de la cuisine. Ensuite, il est utilisé pour générer de nouvelles recettes en fonction de différents critères, tels que les ingrédients spécifiques, les types de cuisine ou les préférences alimentaires. L'évaluation de CookGAN montre qu'il est capable de générer des recettes plausibles et cohérentes, avec des résultats comparables à ceux des recettes réelles. Ils notent également que cette approche peut être utilisée pour générer d'autres types de données culinaires, tels que des images de plats, des descriptions de saveurs, etc

- Lantao Yu, Weinan Zhang .al [14] proposent une approche "SeqGAN : Sequence Generative Adversarial Nets with Policy Gradient" pour la génération de séquences discrètes, telles que des phrases dans le langage naturel, en utilisant des réseaux de neurones générateurs adverses. L'approche combine des techniques de génération de séquences avec une méthode de renforcement par apprentissage. L'algorithme SeqGAN utilise une politique de génération de séquences, apprise à partir des récompenses obtenues en utilisant un discriminateur pour évaluer la qualité des séquences générées. Les expériences ont montré que SeqGAN est capable de générer des phrases cohérentes et naturelles, et qu'il est supérieur à d'autres approches de génération de séquences basées sur des réseaux de neurones. Cette approche peut avoir des applications dans la génération de texte, la traduction automatique et la reconnaissance de la parole.

- Siyuan Pan, Ling Dai . . .al [15] présente une nouvelle méthode "ChefGAN : Food Image Generation from Recipes" qui est l'inverse de la précédente elle consiste à générer des images de nourriture à partir de recettes de cuisine en utilisant un modèle de réseau de neurones appelé ChefGAN. Le modèle prend en entrée une recette et génère une image de nourriture correspondante. Ils expliquent comment ils ont entraîné le modèle en utilisant une grande base de données de recettes et d'images de nourriture pour améliorer la qualité des images générées. Ils ont également évalué leur modèle en comparant les images générées avec les images réelles de nourriture et en demandant à des humains de noter leur qualité visuelle. Les résultats montrent que ChefGAN est capable de générer des images de nourriture de haute qualité à partir de recettes de cuisine.

- Bappaditya Mandal; Niladri B. Puhan . . .al [16] décrit une méthode "Deep Convolutional Generative Adversarial Network-Based Food Recognition Using Partially Labeled Data" pour reconnaître les aliments en utilisant un réseau de neurones antagonistes génératifs à convolution profonde (DCGAN) en utilisant des données partiellement étiquetées. Ils ont expliqué comment ils ont entraîné leur modèle en utilisant une base de données de photos alimentaires partiellement étiquetées pour générer de nouvelles images d'aliments similaires à celles de la base de données. Ils ont également utilisé des techniques de transfert de connaissances pour améliorer la performance de leur modèle. Ils ont évalué leur modèle en comparant sa performance avec celle d'autres modèles de reconnaissance d'aliments. Les résultats montrent que leur modèle DCGAN peut atteindre des performances de reconnaissance d'aliments comparables ou supérieures à celles des autres modèles, même avec des données partiellement étiquetées.

Une autre étude a été menée dans le domaine de la cuisine en utilisant un autre algorithme appelé Transformers.

- Helena H. Lee, Ke Shu . . . al [17] ont décrit une méthode "RecipeGPT : Generative

Pre-training Based Cooking Recipe Generation and Evaluation System” de génération automatique de recettes à l’aide de GPT-2 (un modèle de langage pré-entraîné basé sur des transformers) et d’une technique de pré-entraînement générateur. Ils ont collecté des données d’ingrédients et de recettes à partir de sources en ligne, nettoyé les données, les ont préparées pour l’entraînement et ont entraîné le modèle sur ces données. Le modèle a ensuite été évalué sur sa capacité à générer des recettes cohérentes et plausibles. Les résultats ont montré que le modèle était capable de générer des recettes cohérentes et plausibles.

Il existe d’autres domaines où des études ont été menées sur la génération, et qui pourraient servir d’inspiration pour la génération dans le domaine de la cuisine.

- Yizhe Zhang, Zhe Gan..al [18] ont présenté une méthode ”TextGAN : Generating Text via Adversarial Training” pour la génération de texte à l’aide de réseaux de neurones génératifs adverses (GAN). La méthode consiste à entraîner deux réseaux, un générateur et un discriminateur, qui s’affrontent dans un processus d’entraînement itératif. Le générateur prend en entrée un bruit aléatoire et produit un texte, tandis que le discriminateur évalue si le texte est réel ou synthétique. Les deux réseaux sont formés simultanément en utilisant un signal de récompense de type RL (Reinforcement Learning).Les résultats montrent que la méthode proposée peut générer du texte qui ressemble à du texte réel, en se basant sur des corpus de données tels que des articles de journaux, des romans et des dialogues. Les expériences menées montrent également que l’utilisation de RL dans le processus d’entraînement peut améliorer considérablement la qualité du texte généré. La méthode peut avoir des applications potentielles dans la génération automatique de textes, tels que la création de résumés de textes, la traduction automatique, la génération de dialogues, etc

- Donatas Repecka,Vykintas Jauniskis ... al [19] ils ont décrit une méthode ”Expanding functional protein sequence spaces using generative adversarial networks” pour élargir l’espace des séquences protéiques fonctionnelles en utilisant des réseaux de neurones antagonistes génératifs (GAN). Ils ont expliqué comment ils ont entraîné leur modèle GAN en utilisant une base de données de séquences protéiques fonctionnelles pour générer de nouvelles séquences protéiques qui sont similaires aux séquences fonctionnelles existantes mais qui ont des propriétés différentes. Les auteurs ont évalué leur modèle en testant les nouvelles séquences protéiques générées pour leur fonctionnalité et leur stabilité. Les résultats montrent que la méthode proposée peut générer de nouvelles séquences protéiques fonctionnelles avec des propriétés différentes, ce qui pourrait avoir des implications importantes pour la recherche en biologie et en médecine.

- Nathan Killoran, Leo J. Lee ... al [20] ont décrit une méthode ”Generating and designing DNA with deep generative models” pour générer et concevoir de l’ADN à l’aide de modèles génératifs profonds. Ils ont expliqué comment ils ont utilisé des réseaux de neurones antagonistes génératifs (GAN) et des réseaux de neurones récurrents (RNN) pour apprendre à générer des séquences d’ADN similaires à celles des données d’entraînement. Ils ont également utilisé des techniques d’optimisation pour concevoir de nouvelles séquences d’ADN avec des propriétés spécifiques, telles que la stabilité thermique ou la capacité à se lier à des protéines spécifiques. Ils ont évalué leur méthode en comparant les séquences d’ADN générées avec celles des données d’entraînement et en testant les nouvelles séquences

d'ADN conçues pour leur fonctionnalité. Les résultats montrent que les modèles génératifs profonds peuvent être utilisés pour générer et concevoir de l'ADN avec des propriétés spécifiques, ce qui pourrait avoir des implications importantes pour la biologie synthétique et la médecine personnalisée.

- John Weldon, Tomas Ward [21] . . . al ont décrit une méthode de génération de dossiers médicaux synthétiques pour les études de recherche en utilisant les réseaux de neurones. Ils ont utilisé des réseaux de neurones générateurs adverses (GAN) pour générer des données de santé synthétiques qui ont des caractéristiques statistiques similaires à celles des données de santé réelles. Les GAN utilisés dans cette étude sont spécialement conçus pour les données de séries temporelles, comme les données de santé électroniques. Les résultats montrent que les dossiers médicaux synthétiques générés sont capables de capturer les caractéristiques des données réelles, tout en préservant la confidentialité des patients. Les données synthétiques générées peuvent donc être utilisées pour des études de recherche en santé, sans compromettre la vie privée des patients. Cette approche présente également l'avantage de pouvoir générer des données dans des quantités importantes, ce qui peut être utile pour des études nécessitant de grandes quantités de données. En résumé, cette étude montre que l'utilisation de GAN pour la génération de données médicales synthétiques peut être une méthode prometteuse pour la recherche en santé, offrant une alternative aux données réelles et préservant la confidentialité des patients.

- Charitos Charitou, Simo Dragicevic [22] . . . al ont présenté une méthode "GANs for Fraud Detection" de détection de fraudes financières à l'aide de réseaux de neurones générateurs adverses (GAN). Ils proposent d'utiliser un GAN pour générer des exemples de transactions frauduleuses, qui sont ensuite utilisés pour entraîner un modèle de détection de fraudes. Le GAN est entraîné sur un ensemble de données de transactions légitimes afin de générer des exemples de transactions frauduleuses qui ressemblent aux transactions réelles. En utilisant cette approche, les auteurs ont obtenu des résultats prometteurs pour la détection de fraudes sur un ensemble de données de transactions financières.

- Philippe Schwaller, Teodoro Laino. . . al [23] ont présenté une méthode "Neural Reaction Predictor for Organic Synthesis" basée sur les réseaux de neurones pour prédire les réactions chimiques dans la synthèse organique. Ils ont créé un modèle de réseau de neurones capable de prédire les réactions possibles entre des composés organiques en utilisant des données d'entrée telles que la structure moléculaire, la température et les réactifs. Le modèle a été entraîné sur une grande base de données de réactions organiques et a été évalué en termes de précision de la prédiction. Les résultats ont montré que le modèle était capable de prédire les réactions avec une précision élevée et qu'il pourrait être utile pour accélérer le processus de découverte de nouveaux composés en synthèse organique.

- Artur Kadurin, Sergei Nikolenko . . . al [24] ont décrit un nouveau modèle "druGAN : An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico" de réseau de neurones appelé "druGAN" qui peut être utilisé pour générer de nouvelles molécules avec des propriétés moléculaires souhaitées en utilisant des techniques de simulation informatique. Ils expliquent comment ils ont entraîné leur modèle en utilisant une grande base de données de molécules et en utilisant des techniques d'apprentissage profond pour identifier les ca-

ractéristiques importantes des molécules. Ils ont évalué leur modèle en générant de nouvelles molécules avec des propriétés spécifiques et en comparant leurs propriétés moléculaires avec celles des molécules réelles. Les résultats montrent que le modèle peut générer de nouvelles molécules avec des propriétés souhaitées, ce qui pourrait avoir des implications importantes pour la découverte de médicaments et la chimie des matériaux.

Une étude supplémentaire a été réalisée en utilisant un algorithme différent, à savoir l'algorithme des colonies de fourmis.

- Dong Do Duc ; Phue Thai Dinh . . . al[25] ont présenté une méthode efficace "An Efficient Ant Colony Optimization Algorithm for Protein Structure Prediction" présente une méthode pour prédire la structure des protéines en utilisant un algorithme d'optimisation de colonie de fourmis (ACO). Ils expliquent comment ils ont adapté l'ACO pour résoudre le problème de la prédiction de la structure des protéines en utilisant une représentation discrète de la structure des protéines et en combinant des heuristiques spécifiques au domaine avec l'ACO. Ils ont évalué leur méthode en utilisant des benchmarks standard pour la prédiction de la structure des protéines et ont comparé leurs résultats avec ceux d'autres méthodes de prédiction de la structure des protéines. Les résultats montrent que l'ACO peut être une méthode efficace pour prédire la structure des protéines avec une précision comparable à celle des autres méthodes.

Une autre étude a été menée en utilisant un algorithme différent, à savoir l'algorithme des colonies d'abeilles

- Ammar K Alazzawi, Helmi Md Rais et Shuib Basri [26] ont présenté une nouvelle méthode "ABCVS : An Artificial Bee Colony for Generating Variable T-Way Test Sets" pour générer des ensembles de tests pour les tests de combinaison, appelés "test sets T-way". Les tests T-way sont utilisés pour tester les interactions entre les différentes caractéristiques d'un système, ce qui peut aider à détecter des erreurs dans le système. La méthode proposée, ABCVS (Artificial Bee Colony for Generating Variable T-way Test Sets), est basée sur l'algorithme d'optimisation ABC (Artificial Bee Colony). ABCVS utilise une approche multi-objectifs pour générer des ensembles de tests T-way optimaux qui ont une couverture maximale des interactions entre les caractéristiques.

Dans la cuisine saoudienne des recherches ont été faites pour explorer l'hypothèse du food pairing en utilisant les algorithmes génétiques. Les auteurs ont développé un modèle quantitatif et l'ont appliqué à un ensemble de données composé de composés d'ingrédients trouvés dans les recettes de la cuisine saoudite[3]. Un autre travail réalisé par Kerbedj et al. [27], au sein de notre laboratoire, a également exploré les algorithmes génétiques pour étudier de l'hypothèse de food pairing dans la cuisine algérienne. Cette technique de résolution de problèmes d'optimisation avec ou sans contraintes, basée sur la sélection naturelle, qui régit l'évolution biologique a été employée pour générer l'ensemble des recettes synthétiques.

Dans la continuité de ce travail, nous nous intéressons à l'étude de l'hypothèse de food pairing dans la cuisine maghrébine. Nous explorons le modèle LSTM (Long-Short-Term-Memory) pour la génération des recettes synthétiques nécessaire au calcul de la tendance d'une cuisine dans l'appariement d'ingrédients qui constituent ces recettes. Le

modèle LSTM est une forme de réseau de neurones récurrents (RNN) utilisée en apprentissage profond (deep learning). Contrairement aux RNNs, les LSTMs ont des connexions récurrentes qui leur permettent de traiter des séquences de données, tels que des textes, des signaux audio ou vidéo. Grâce à cette fonctionnalité, les LSTMs sont utilisés pour diverses tâches telles que la reconnaissance de l'écriture manuscrite continue ou la reconnaissance vocale [28].

Enfin, après avoir examiné un certain nombre d'approches, le modèle LSTM semble prometteurs pour donner de bons résultats grâce à sa capacité de génération de séquences textuelles. La capacité des LSTMs à modéliser efficacement diverses séquences la distingue des autres architectures. Cette méthode promet d'apporter une nouvelle couche à la génération de recettes synthétiques en offrant de la variété, de la créativité et une meilleure qualité.

1.5 Conclusion

Dans ce chapitre, nous avons présenté que la cuisine est un langage universel qui utilise le food pairing pour créer des combinaisons savoureuses à partir de composés aromatiques communs, et nous avons mentionné certains travaux existants à propos de l'hypothèse de food pairing et de certains approches réalisées sur la génération , et enfin on a montré ce qu'on a choisi comme approche pour résoudre le problème qu'on a mentionné.

Chapitre 2

Approche proposée

2.1 Introduction

Dans ce chapitre, nous examinerons l'hypothèse de food pairing pour voir s'il existe des tendances positives ou négatives dans la cuisine maghrébine. Plus précisément, nous verrons si les recettes maghrébines préfèrent utiliser des ingrédients dont les composantes gustatives sont similaires ou pas. Nous commencerons par présenter le Schéma global de la solution que nous proposons. Ensuite, nous décrirons les procédures et les méthodes utilisées pour mener à bien toutes ces tâches.

2.2 Schéma global

Notre méthode comporte quatre étapes principales pour atteindre cet objectif : collecte et prétraitement de l'ensemble de données ; utilisation de l'architecture LSTM pour générer des recettes aléatoires ; calcul de la moyenne des saveurs partagées entre chacune des recettes réelles et générées ; et enfin, comparaison des deux cuisines pour déterminer la tendance de la cuisine maghrébine. La technique proposée est résumée dans la figure 3.1. La figure 2.1 présente le résumé de haut niveau de la proposition.

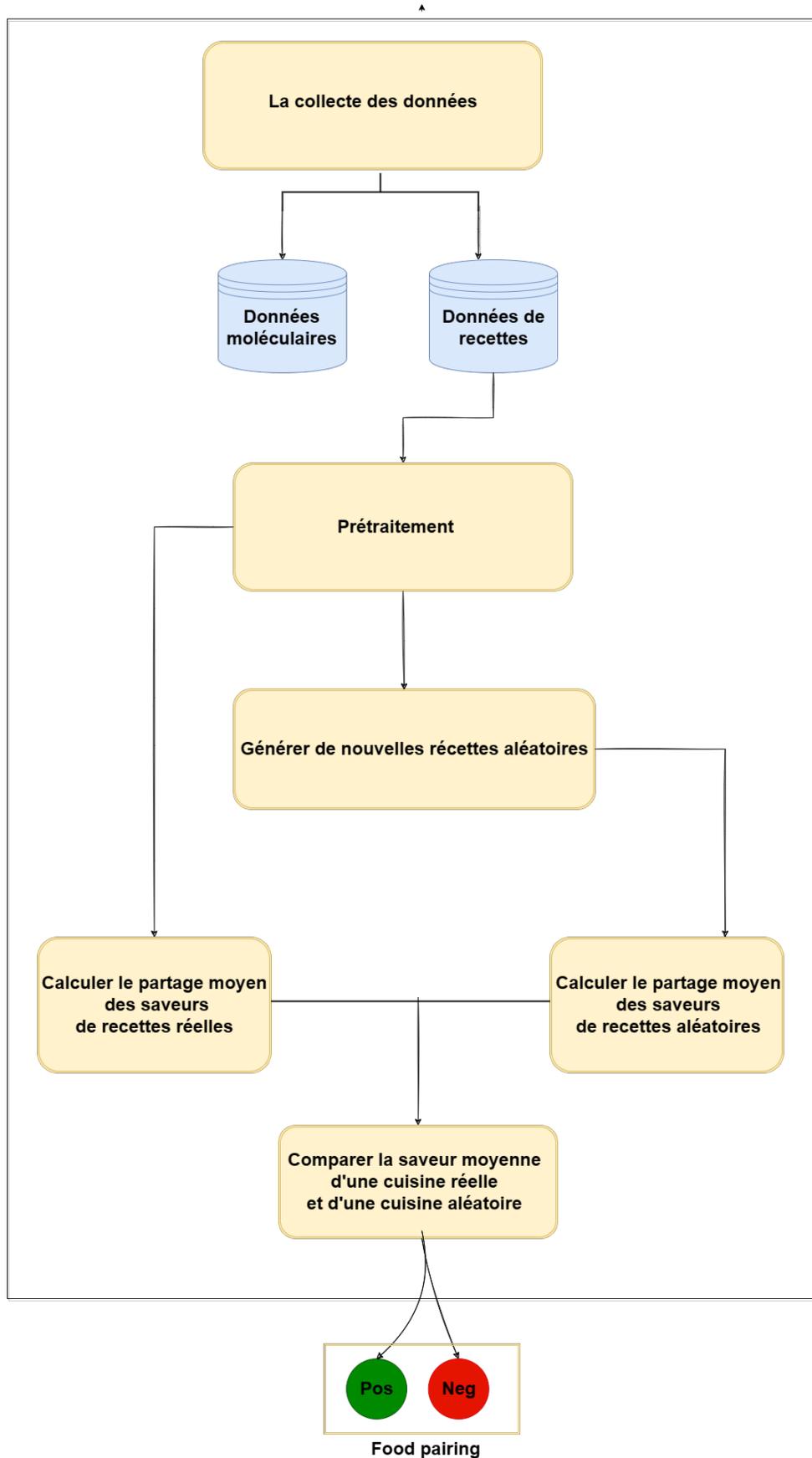


FIGURE 2.1 – Aperçu¹⁹ de l'approche proposée

2.3 La collecte des données

Nous avons utilisé deux ensembles de données qui sont nécessaires pour notre projet : le premier est le livre de cuisine traditionnelle écrit par Salima HADJIAT[29] afin de développer notre propre collection de créations culinaires. Nous avons choisi ce livre parce qu'il ne contient que des recettes traditionnelles qui existent depuis longtemps. Le deuxième est la base de données moléculaire [9] qui a été présenté plus haut.

2.4 Prétraitement

Le pré traitement des données, ou le processus de conversion des données brutes en un format clair et structuré, est une étape importante car la qualité des données utilisées a un impact direct sur le fonctionnement des algorithmes. Dans notre cas, un livre de cuisine qui comprend des recettes, une liste d'ingrédients avec la quantité de chacun d'entre eux à utiliser, ainsi que des instructions de préparation. L'objectif de cette opération est d'encoder chaque recette à partir d'un format de livre brut avec ses ingrédients uniques dans un ensemble de données ordonné et organisé. Pour pré traiter nos données, nous avons suivi trois étapes principales : le nettoyage des données, l'intégration des données et la transformation des données.

2.4.1 Nettoyage des données

Il s'agit d'une méthode permettant d'éliminer les données imprécises, insuffisantes ou erronées d'un ensemble de données.

2.4.2 l'intégration des données

À ce stade, nous sollicitons de nombreuses sources pour produire un ensemble de données assez important qui comprend autant de données que possible sur la question. Par exemple, nous avons utilisé cette méthode pour notre base de données moléculaires ; la majorité des informations avaient déjà été extraites d'internet et sauvegardées dans un fichier, mais lorsque nous avons constaté que certains composants manquaient, nous avons dû rechercher d'autres sources dans la littérature scientifique afin d'ajouter leurs profils gustatifs à notre ensemble de données.

2.4.3 Transformation des données

Dans cette phase, nous transformons les données dans un format approprié pour les étapes suivantes. Pour nos données, soit nous remplaçons les ingrédients manquants dans

notre base par d'autres ingrédients pertinents ou bien nous supprimons d'autres ingrédients manquants avec leurs recettes si ces ingrédients sont essentiels et si ce n'est pas le cas nous supprimons complètement les ingrédients. comme présenté à la figure 2.2.

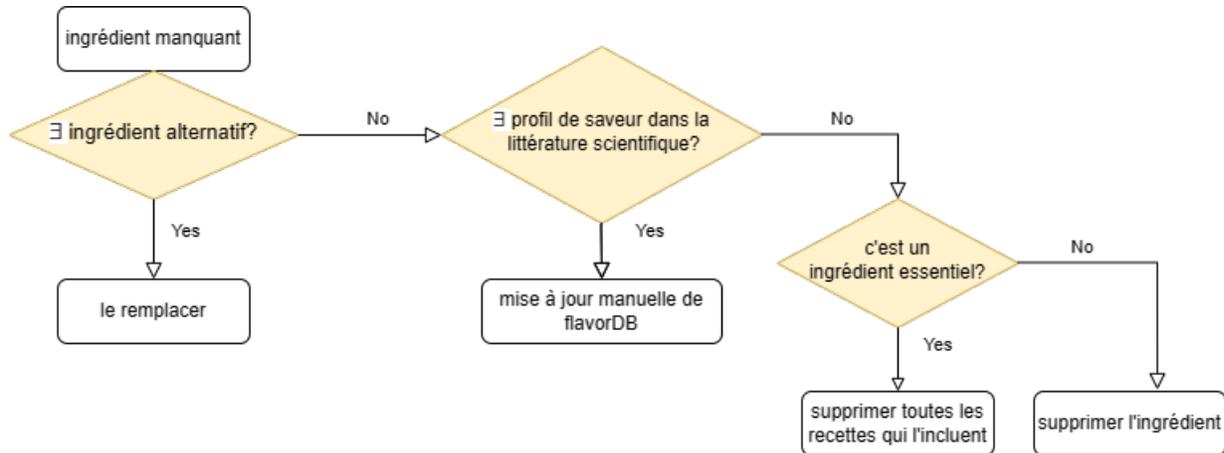


FIGURE 2.2 – Diagramme illustrant le prétraitement des données

2.5 Génération de recettes aléatoires

Nous avons utilisé le test décrit dans l'une des premières recherches dans ce domaine, menée par Yong-Yeol Ahn et ses collègues[2], pour effectuer l'analyse de food pairing. Ce test a ensuite été utilisé dans tous les articles ultérieurs[3], [10] et [11]. Si l'hypothèse est correcte, l'ensemble des plats de la cuisine étudiée devrait présenter en moyenne plus d'éléments gustatifs communs que n'importe quel autre groupe de recettes. La première étape de ce test consiste à créer un groupe de recettes choisi suivant une méthode qui sera décrite par la suite, cet ensemble sera dans un second temps comparé à l'ensemble réel.

2.5.1 Le processus de génération

Un aspect essentiel de la vérification de l'hypothèse est la création des recettes. Pour ce faire, nous utilisons l'architecture LSTM. Pour bien comprendre notre processus de génération on doit d'abord comprendre son architecture[30]. Un réseau LSTM typique est composé de cellules, qui sont les formes rectangulaires illustrées dans la figure ci-dessous. L'état de la cellule et l'état caché sont les deux états qui sont transmis à la cellule suivante. Les trois principaux processus, appelés gates, sont utilisés pour manipuler les blocs de mémoire, qui sont responsables de la mémoire. La discussion de chacun d'entre eux est présentée ci-dessous.

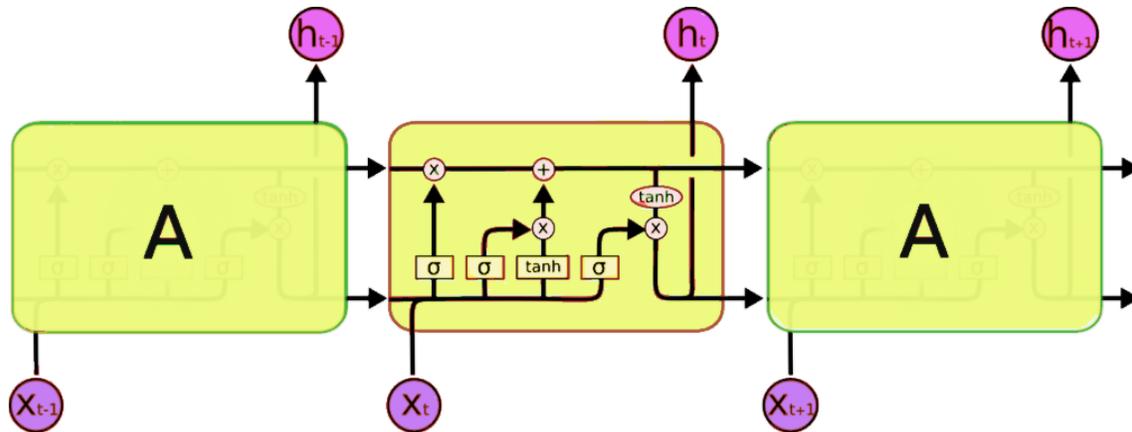


FIGURE 2.3 – Figure illustrant l'architecture du lstm

Porte d'oubli :

L'information provenant de l'état de la cellule est éliminée par une porte d'oubli. La multiplication d'un filtre permet d'éliminer les informations qui ne sont plus nécessaires à l'appréhension des choses par la LSTM ou qui sont moins pertinentes. Ceci est nécessaire pour améliorer les performances du réseau LSTM.

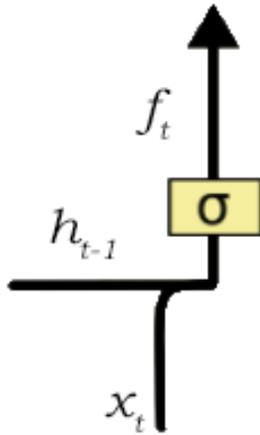


FIGURE 2.4 – Porte d’oubli

h_{t-1} et x_t sont les deux entrées de cette porte. L’entrée à ce pas de temps spécifique est x_t , tandis que l’état caché de la cellule précédente, ou sa sortie, est h_{t-1} . Les matrices de poids sont multipliées par les entrées avant qu’un biais ne soit introduit. La fonction sigmoïde est ensuite appliquée à cette valeur. Chaque nombre dans l’état de la cellule est représenté par un vecteur dont les valeurs sont comprises entre 0 et 1 à la suite de l’application de la fonction sigmoïde. En substance, la fonction sigmoïde détermine les valeurs à conserver et celles à rejeter. Lorsque la porte d’oubli émet un "0" pour une valeur spécifique de l’état de la cellule, cela signifie que la porte d’oubli veut que l’état de la cellule oublie complètement cet élément d’information particulier. Un "1" indique que la porte d’oubli souhaite rappeler l’ensemble des données, de la même manière. L’état de la cellule est multiplié par la sortie de ce vecteur de la fonction sigmoïde.

Porte d'entrée :

La porte d'entrée est chargée d'ajouter de l'information à l'état de la cellule. Le diagramme ci-joint illustre les trois étapes qui constituent le processus fondamental de cet ajout d'informations.

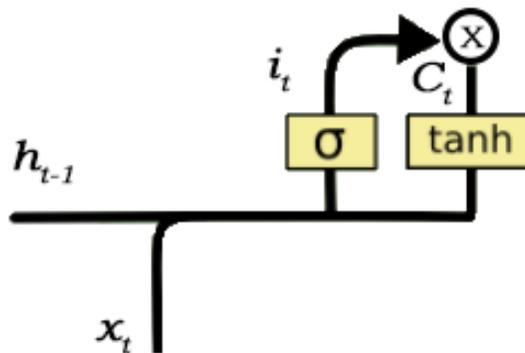


FIGURE 2.5 – Porte d'entrée

-La fonction sigmoïde contrôle les valeurs à ajouter à l'état de la cellule. Elle fonctionne comme un filtre pour toutes les données provenant de h_{t-1} et x_t et est fondamentalement très similaire à la porte d'oubli.

-Créer un vecteur contenant chaque valeur susceptible d'être ajoutée à l'état de la cellule (tel que déterminé par h_{t-1} et x_t). La fonction \tanh , qui produit des valeurs comprises entre -1 et +1, est utilisée à cet effet.

-Ajouter ces données précieuses à l'état de la cellule par une opération d'addition après avoir multiplié la valeur du filtre régulateur (la porte sigmoïde) par le vecteur produit (la fonction \tanh). Une fois cette procédure en trois étapes terminée, nous nous assurons que seules les informations les plus cruciales et non redondantes sont ajoutées à l'état de la cellule.

Porte de sortie :

Là encore, le fonctionnement d'une porte de sortie peut être divisé en trois étapes :

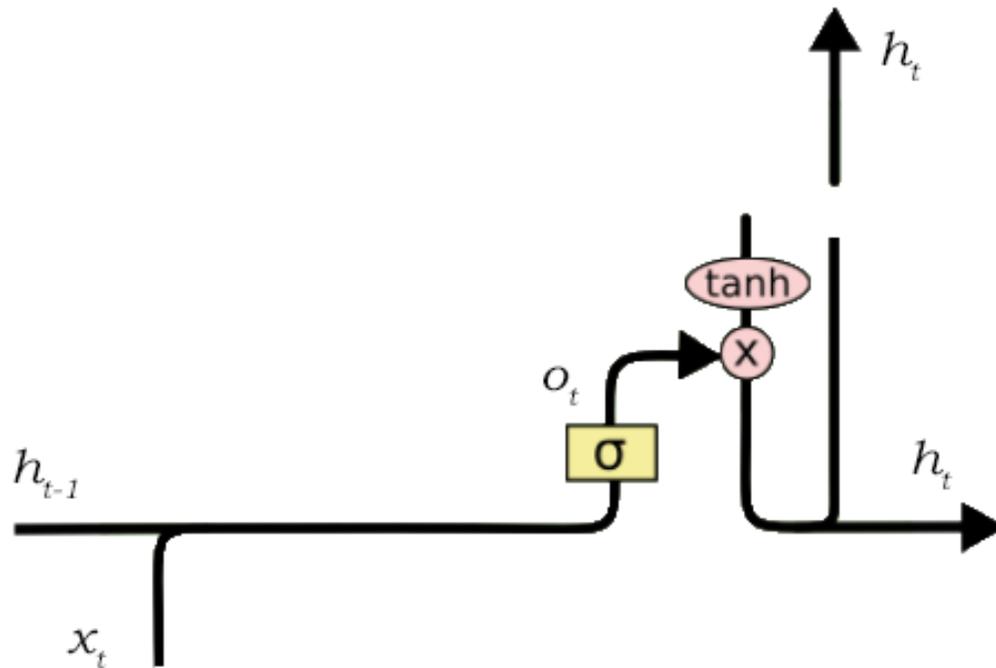


FIGURE 2.6 – Porte de sortie

-Après la mise à l'échelle des valeurs dans la plage de -1 à +1 à l'aide de la fonction tanh, un vecteur est produit.

-Création d'un filtre capable de contrôler les valeurs qui doivent être produites à partir du vecteur établi ci-dessus en utilisant les valeurs h_{t-1} et x_t . La fonction sigmoïde est à nouveau utilisée dans ce filtre.

-Ajouter la valeur de ce filtre régulateur au vecteur produit à l'étape 1, la multiplier et envoyer le résultat à la fois à la sortie et à l'état caché de la cellule suivante.

2.6 Calcul du partage moyen des saveurs

La moyenne des composants chimiques du goût qui sont partagés par tous les ingrédients des recettes est estimée par le partage moyen des saveurs d'une cuisine donnée. Pour ce faire, nous calculons d'abord la même mesure pour chaque recette, puis la moyenne pour l'ensemble du menu. Ces procédures s'appliquent à la fois aux recettes réelles et fictives.

Cette procédure s'applique à la fois aux aliments réels et fictifs. Les annotations suivantes seront utilisées pour formaliser cette procédure : La cardinalité de l'intersection des ensembles $F(i)$ et $F(j)$, qui est le nombre de composés aromatiques partagés par deux composants (i, j) , peut être calculée à l'aide de l'équation suivante.

$$Ns(i, j) = |F(i) \cap F(j)| \quad (2.1)$$

Pour obtenir le partage moyen des saveurs, cette formule est appliquée à toutes les combinaisons des composants de la recette R et introduite dans l'équation $Ns(R)$.

$$Ns(R) = \frac{2}{nR(nR - 1)} \sum_{i,j \in R; i \neq j} |F(i) \cap F(j)| \quad (2.2)$$

Comme $Ns(R)$ représente le partage moyen des saveurs pour une recette, il faut répéter la procédure pour chaque recette de la cuisine afin d'obtenir le partage moyen des saveurs de l'ensemble de la cuisine, comme le montre la formule ci-dessous.

$$Ns = \sum_R \frac{Ns(R)}{Nc} \quad (2.3)$$

Où :

Symbole	Description
R	représente une recette
$Ns(i, j)$	nombre de composés partagés entre les les ingrédients i et j
$F(i)$	Ensemble de composés aromatiques trouvés dans un ingrédient i
$Ns(R)$	partage moyen des saveurs de la recette R
nR	nombre d'ingrédients dans la recette R
Ns	Saveur moyenne de la cuisine
Nc	Nombre de recettes dans la cuisine

TABLE 2.1 – Annotations des formules ci-dessus et leur description

2.7 Tester l'hypothèse de food pairing

Nous disposons déjà de toutes les informations nécessaires pour valider ou rejeter la théorie susmentionnée, à savoir si les aliments souvent utilisés dans la cuisine maghrébine contiennent des composantes gustatives similaires ou, au contraire, Pour ce faire, nous devons déterminer la différence entre le score de partage moyen des saveurs de la cuisine maghrébine authentique "Ns(real)" et le score de partage moyen des saveurs des recettes aléatoires "Ns(rand)" à l'aide de l'équation suivante :

$$\Delta N = Ns(real) - Ns(rand) \quad (2.4)$$

Nous pouvons étudier les tendances de la cuisine maghrébine à l'aide des résultats de cette méthode. Si $\Delta N > 0$, un couple d'aliments est positif et par conséquent ,l'hypothèse de food pairing est positive, sinon il est négatif.

2.8 Conclusion

Dans ce chapitre, nous avons examiné la procédure et les étapes générales que nous avons suivi pour mettre notre stratégie en pratique. Nous avons commencé par un aperçu général de notre solution, puis nous avons décomposé les tactiques utilisées à chaque étape.

Chapitre 3

Test et validation

3.1 Introduction

Nous avons procédé à de nombreuses étapes de préparation des données afin de mener à bien notre recherche, en commençant par la collecte des recettes maghrébines traditionnelles, suivie de l'extraction des composants spécifiques. Nous avons aussi relié ces composés à leurs parties moléculaires afin de produire un caractère gustatif. Dans ce chapitre, nous verrons comment notre stratégie a été mise en pratique, ainsi que tous les processus précédents. Enfin, nous résumons l'expérience que nous avons menée et les résultats obtenus.

3.2 Environnement de travail :

Nous présenterons ici notre environnement de travail, ainsi que les programmes et bibliothèques que nous avons utilisés : Langage de programmation : Python 3 Bibliothèques utilisées : re, pattern, NLTK, pandas, enchant, deep translator, wordcloud, plotly, NumPy

3.2.1 Python

Python est un langage de programmation très apprécié. En 1991, Guido van Rossum l'a créé et il est devenu disponible. Il est utilisé pour : développement web (côté serveur), développement de logiciels, mathématiques et scripting de système. Python est utilisé comme langage de programmation principal pour développer et implémenter le code de notre projet de recherche. [31]

3.2.2 Bibliothèque d'expressions régulières

La majorité des travaux de traitement de texte s'appuient sur les expressions régulières (re). Re, le module d'expressions régulières intégré à Python, a été notre dernier choix. Les

chaînes de 8 bits (bytes) et les chaînes Unicode (str) sont toutes deux supportées[32]. et dans notre cas, les fonctions fournies par défaut se sont avérées adéquates.

3.2.3 Bibliothèque des Pandas

Pandas est un outil open-source rapide et flexible pour la manipulation de données[33]. Un cadre de données, une représentation tabulaire des données, est la principale structure de données utilisée par pandas. Le code critique du dataframe est écrit en C ou en Python et est largement optimisé pour la vitesse. La bibliothèque des Pandas a été utilisée dans le code pour lire et manipuler les données des fichiers CSV, , fournir des fonctionnalités puissantes pour effectuer des opérations de nettoyage et de transformation des données, telles que la lecture des données en tant que DataFrame, l'extraction de listes d'ingrédients.

3.2.4 Bibliothèque NLTK

Connue également sous le nom de Natural Language Toolkit, cette bibliothèque Python commune à code source ouvert a été créée par Edward Loper et Steven Bird du département d'informatique et des sciences de l'information de l'université de Pennsylvanie[34]. Il offre un large éventail de fonctionnalités NLP, notamment l'analyse syntaxique, la lemmatisation, l'étiquetage et l'extraction. NLTK a été utilisé pour améliorer la représentation des ingrédients et des recettes, ainsi que pour faciliter l'analyse linguistique dans le cadre de notre projet de génération de recettes et d'analyse des saveurs.

3.2.5 Bibliothèque Deep translator

Une bibliothèque tierce à code source ouvert appelée Deep translator prend en charge un grand nombre de fournisseurs de traduction, notamment Google Translate, Yandex et Microsoft Translator. De nombreuses fonctionnalités, dont la traduction par lots et l'identification automatique de la langue, sont prises en charge dès le départ.

3.3 Bibliothèque Wordcloud

Wordcloud est une application de visualisation qui permet de créer des nuages de mots entièrement personnalisables, y compris la prise en charge des mots d'arrêt, la couleur, la taille, etc. Wordcloud a été utilisée pour générer des visualisations de nuages de mots à la fois avant et après le traitement des données

3.4 Bibliothèque Plotly

Plotly de Python est une boîte à outils graphique interactive et open-source basée sur le web[35]. Il dispose déjà de graphiques en 3D, d'animations et d'une multitude d'options de

personnalisation des tracés. Plotly a été utilisée pour créer des graphiques, des diagrammes et des visualisations pour présenter les résultats

3.5 Bibliothèque NumPy

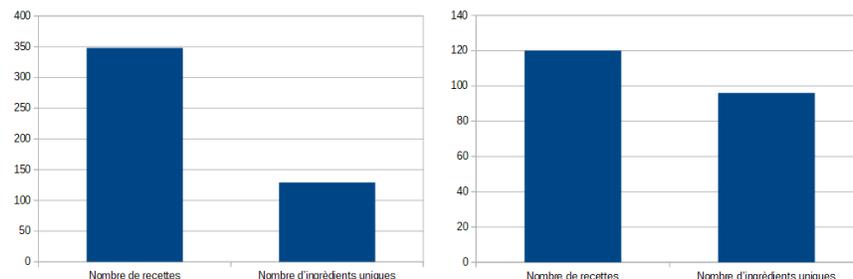
L'acronyme NumPy signifie en réalité "Numerical Python". Il s'agit d'une bibliothèque Python open-source. Cet outil est utilisé pour la programmation scientifique en Python, en particulier pour la programmation dans les domaines de la science des données, de l'ingénierie, des mathématiques ou des sciences. La bibliothèque NumPy a été utilisée dans le code pour effectuer des opérations numériques et de manipulation de tableaux multidimensionnels

3.6 Ensemble de données :

3.6.1 Recettes traditionnelles maghrébines

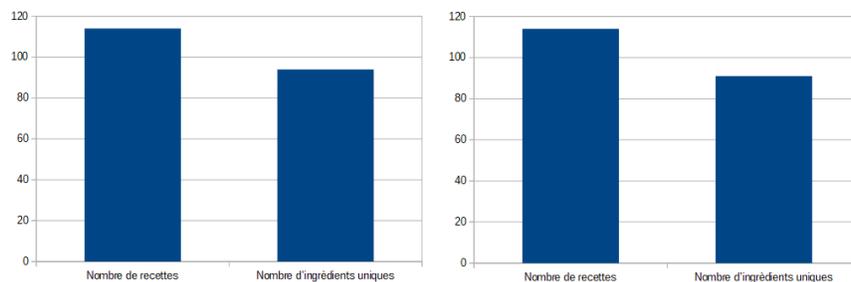
Le célèbre livre de cuisine maghrébine de Mme Salima HADJIAT est celui que nous avons utilisé pour obtenir les données. Il comprend plus de 400 plats maghrébins authentiques provenant de divers pays : l'Algérie, le Maroc et la Tunisie, ainsi que des recettes de poissons, de légumes, de soupes, de tajine sucré et non sucré, d'entremets et d'autres catégories de cuisine. Mais en raison de son ancienneté il n'existe pas une version pdf ou une version en ligne. Nous avons donc dû écrire manuellement le contenu avant de procéder au véritable pré traitement.

Les figures 3.1 et 3.2 montrent le nombre de recettes et le nombre d'ingrédients unique dans chaque cuisine.



(a) Nombre de recettes et (b) Nombre de recettes et d'ingrédients uniques dans la cuisine maghrébine d'ingrédients uniques dans la cuisine algérienne

FIGURE 3.1 – Diagramme de nombre de recettes et d'ingrédients uniques pour la cuisine maghrébine et algérienne



(a) Nombre de recettes et (b) Nombre de recettes et d'ingrédients uniques dans la cuisine tunisienne et marocaine

FIGURE 3.2 – Diagramme de nombre de recettes et d'ingrédients uniques pour la cuisine tunisienne et marocaine

3.6.2 Ensemble de données moléculaires :

La base de données moléculaires que nous avons utilisée pour relier le profil gustatif de chaque ingrédient à son arôme est "FlavoDB" [9]. Il s'agit d'une collection de caractéristiques moléculaires. Cette base de données est la mieux adaptée à notre recherche car elle met l'accent sur la composante chimique des goûts et donne des informations sur les endroits où l'on peut se procurer les produits chimiques de manière naturelle [9]. La base a été créée pour rassembler les informations les plus complètes sur les molécules gustatives dans une seule base de données. Pour ce faire, toutes les données essentielles ont été recueillies auprès de diverses sources. La base de données moléculaire contient 25595 molécules gustatives, dont 2254 sont liées à 936 composants [9].

3.7 Pré traitement

La première étape a consisté à séparer et à préserver la liste des ingrédients de la partie de la recette consacrée à la préparation du repas. Pour ce faire, nous avons utilisé une expression régulière qui délimite les recettes avec l'expression "composition" comme mot-clé de départ et n'importe quel nombre (d+) comme mot-clé de fin.

3.7.1 Nettoyage des recettes :

Catégories non pertinentes éliminées

Certaines parties du livre ont été immédiatement écartées car elles n'apportent pas des informations pertinentes sur le plan gastronomique :

- La rubrique "Conserves", qui comprend un total de 16 recettes expliquant comment conserver les légumes, les fruits et de viande. Cependant, il suffit de quelques éléments, comme le sel et l'eau, pour préparer de la même manière plusieurs produits.

- La rubrique "Confitures", qui comprend un total de 16 recettes aussi expliquant comment préparer du confiture. Cependant, il suffit de quelques éléments, comme le sucre et l'eau, pour préparer de la même manière une confiture.

Mots supplémentaires

Lorsque nous avons obtenu la liste des composants de chaque plat, nous avons constaté qu'elle comprenait également des termes de mesure (tels que 1 kg, une poignée, amère, etc.) et des adjectifs descriptifs. Nous avons finalement supprimé ces mots, tout en veillant à ne pas en faire trop. éliminer les éléments composés de nombreux mots, par exemple "pomme de terre, huile d'olive...".

Ingrédients multiples dans la même ligne

Il existe deux versions de ce problème où de nombreux composants sont spécifiés dans la même ligne :

-Tout d'abord, lorsque la recette proposait des ingrédients de substitution, tels que "1 cuillerée de beurre ou de smen", nous avons choisi de ne conserver que la première recommandation afin de préserver les composants originaux de la recette.

-La seconde était lorsque des éléments significatifs figuraient sur la même ligne mais étaient séparés par la conjonction "et", comme dans "1 verre d'huile et de beurre". Nous avons remplacé le "et" par "-" pour considérer le deuxième ingrédient comme un élément distinct et indépendant, car les parties en question ne sont pas facultatives.

Ingrédients composés

L'un des problèmes les plus difficiles auxquels nous avons été confrontés concernait les ingrédients composés propres à la cuisine maghrébine, tels que "Merguez, Rechta, Crème simple...etc", plus les ingrédients composés qui ont déjà été trouvés dans l'étude de la cuisine algérienne "Plombs, Tlitli, Crêpes, dioul...etc". Ces produits, qui sont souvent inclus comme ingrédients dans les recettes, sont en réalité des combinaisons de plusieurs ingrédients fondamentaux qui sont énumérés dans une autre recette. Nous avons également remplacé les composants combinés par leurs constituants. Le tableau 3.1 énumère les composants qui constituent le mélange ainsi qu'un résumé des éléments constitutifs de chaque ingrédient. En outre, nous incluons les épices composées qui ont été remplacées dans le tableau 3.2, ainsi qu'une ventilation de leur composition.

Nom de la recette des ingrédients composés supprimé	Liste des éléments de la recette
Crème simple	lait, œufs, maïzena, sucre, eau de rose
Trid/Malsouqa /Bastèla	semouline, farine, sel, eau
Ftayer	farine, vinaigre, pain, sel, eau, huile
Maassems aux amandes	semouline, sel, huile, amandes, sucre, eau de rose, eau, farine, jus de citron
Fruits secs	noix, noisettes, pistache, amande
Crème blanche	lait, sucre, maïzena, eau de fleur oranger
Kefta	viande, oignon, persil, sel, poivron noir, piment fort
Mhammsa	eau, semouline, sel
Sauce tomate	tomate, beurre, sel, poivre noir, persil, thym, laurier, ail, sucre, oignon
Levure de boulanger	farine, vinaigre, pain, sel, eau
Rechta	semouline, sel, huile, farine
Vol au vent	farine, vinaigre, eau, sel, beurre
Sucre vanillé	sucre, vanille
tous les légumes de la saison	pomme de terre, carotte, navet, courgettes, poireau
bouquet garni	carottes, navet, poireau, céleri
Merguez	viande, ail, poivre rouge, cumin, cannelle, coriandre, allspice, clove, gingembre, poivre, piment, tabel, menthe, graines de fenouil, poivre noir, sel, huile

TABLE 3.1 – Liste des recettes remplacées par la liste de leurs ingrédients

Nom des épices composées supprimées	Liste des composants de l'ingrédient de base
Dersa	ail, cannelle, cumin, piment, sel, laurier
Ras lhanout	cumin, gingembre, sel, poivre, cannelle, coriandre, allspice, clove, piment
Chermoula	oignon, ail, persil, poivre rouge, cumin, poivre noir, citron, farine, sel, huile
Hror	poivre, cannelle, clou de girofle, coriandre, noix de muscade, gingembre, cardamome

TABLE 3.2 – Liste des épices composées remplacées par leurs compositions

Ingrédients synonymes

Nous avons découvert plusieurs aspects qui étaient mentionnés différemment dans d'autres sections, et nous avons donc unifié ces synonymes en sélectionnant un terme

Problèmes de traduction

Au cours du processus de traduction automatique, nous avons rencontré quelques problèmes de traduction erronée. Tout d'abord, un certain nombre d'ingrédients différents avaient la même traduction, ce qui était gênant car cela permettait de relier un élément au mauvais profil gustatif d'un ingrédient. Par exemple, "piment rouge" et "poivre rouge" signifient tous deux "red pepper", mais lorsque nous consultons la base de données, nous constatons que le "red pepper" n'est même pas mentionné. Nous avons donc cherché d'autres noms pour les mêmes substances, et le terme "piment rouge" a été remplacé par "capsicum".

Les noms des composants dans notre base de données moléculaire ne correspondaient parfois pas aux résultats de la traduction, ce qui constituait le deuxième problème. La base de données des molécules de saveur stocke généralement les ingrédients sous forme solitaire, alors que les traductions de certains composants étaient au pluriel. Nous avons utilisé la fonction `singularize()` du module d'exploration du web `pattern.en` pour renvoyer la version singulière des noms pluriels, ainsi que la bibliothèque `Enchant` pour vérifier l'orthographe et réparer les fautes d'orthographe probables.

Ingrédients manquants

Malgré la grande variété des 936 éléments de la base de données moléculaire, plus les ingrédients qui ont été ajoutés lors de l'étude de la cuisine algérienne (l'huile d'olive et la levure chimique...etc), nous avons constaté que plusieurs ingrédients, tels que le jaune d'oeuf, le blanc d'oeuf et les grains d'anise vert, manquaient lorsque nous avons associé la base de données de recettes à la base moléculaire. Nous avons remédié à ce problème de plusieurs manières, en tenant compte de chaque situation séparément :

Ajout du profil de saveur : Afin d'éviter autant que possible de modifier la base de données des recettes, nous avons d'abord utilisé cette stratégie chaque fois que cela était possible. Pour ce faire, nous avons recherché une liste de molécules aromatiques qui décrivent le mieux l'article dans d'autres sources. Cette stratégie est efficace. Nous avons également utilisé le jaune d'oeuf, que nous avons été autorisés à ajouter à la suite de l'étude de Akshita Mehta et Reena Gupta [36]. Le blanc d'oeuf suite de l'étude de l'article *Biology of eggs* [37], et nous avons été autorisés aussi à ajouter aussi les grains d'anis vert à la suite de l'étude de [38].

Remplacement d'un ingrédient par un substitut approprié : Nous avons décidé de substituer d'autres composants de la base moléculaires à certains composants manquants, tels que "smen" pour "butter" et "Biscotte" pour "bread" et "Intestin, Tripe, Langue, Rognons, Boyau, Foie, Tête d'agneau ou de mouton, Cerveau" pour "meat".

Supprimer l'ingrédient de la recette : Nous avons décidé d'éliminer les éléments inutiles d'un plat et de nous en tenir à l'original. Cette méthode a été utilisée, par exemple, lorsque le "gras", "absinthe" et "variantes" étaient des composants qui n'étaient pas toujours nécessaires.

Nom de la recette
Briouats à la viande hachée
Rate de veau farcie
Brochettes de kouah
Brochettes de Kefta
Brochettes de foie de veau
Brochettes de Kebab
Brochettes Tunisiennes
Merguez
Bekbouka tunisienne
Abats en "chminka"

TABLE 3.3 – Liste des recettes où la matière grasse est facultative

Nom de la recette
Fricassés

TABLE 3.4 – Liste de recette où l'ingrédient variantes est facultative

Nom de la recette
Escargots

TABLE 3.5 – Liste de recette où l'ingrédient absinthe est facultative

3.8 Visualisation

En conclusion, nous avons commencé avec plus de 400 recettes du livre de la cuisine, nous en avons éliminé plusieurs pour les raisons évoquées précédemment dans ce chapitre , et nous avons finalement obtenu 348 recettes avec 129 ingrédients uniques pour notre base de données avec une moyenne de 11 ingrédients par recette, un minimum de 4 ingrédients et avec un maximum de 41 ingrédients. La figure 3.5 montre les fréquences des ingrédients dans la cuisine maghrébine.

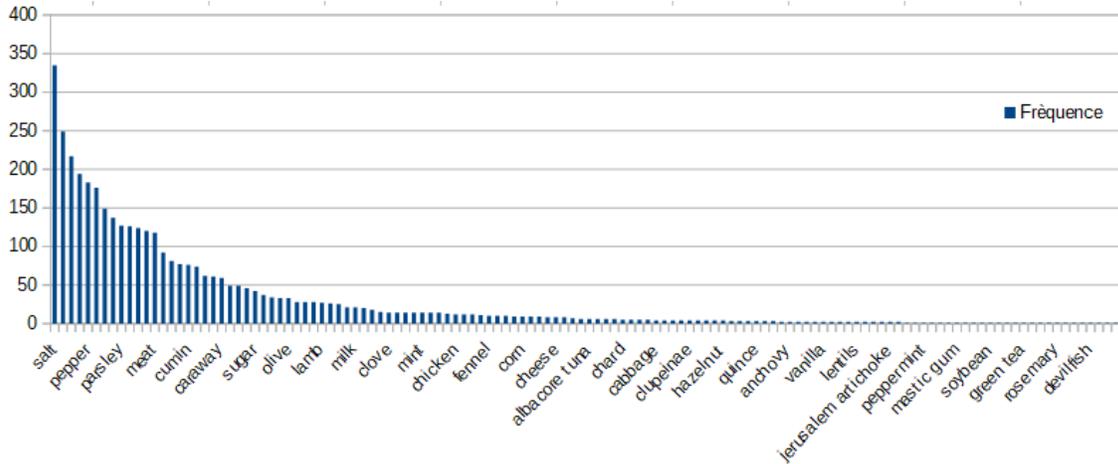


FIGURE 3.5 – La fréquence des ingrédients dans la cuisine maghrébine

3.9 Processus de génération

Dans la partie suivante, nous créerons des recettes aléatoires à l'aide de l'architecture LSTM, algorithmes génétiques et génération aléatoire puis nous testerons l'hypothèse de Food pairing à l'aide de l'équation 2.4 $\delta Ns = Ns(real) - Ns(rand)$ qui calcul le partage moyen des saveurs entre les recettes réelles et celles générées par notre algorithme. Après les avoir identifiées et comparées pour obtenir une conclusion sur le food pairing dans la cuisine maghrébine, nous avons évalué plusieurs configurations pour différents facteurs.

3.9.1 Génération de recettes synthétiques

Ensuite, nous avons utilisé trois méthode de génération de recettes synthétiques : les algorithmes génétiques, génération aléatoire et l'agorithme lstm, avec la condition d'arrêt suivante pour générer un ensemble de recettes aléatoires :

Toutes les approches reposent sur une condition d'arrêt qui plafonne la quantité de recettes nouvellement développées relative à la taille de notre ensemble de données initial. Pour chaque test de génération, le processus a été exécuté 100 fois. En faisant varier les paramètres de chaque méthode.

- Paramètres de GA étudiés :
 - Taille de la recette : Nous avons déterminé la taille de la recette à générer (pour l'algorithme génétique) en fonction de la taille moyenne des recettes de notre ensemble de données. Par conséquent, nous avons fixé la taille à 11 pour l'ensemble de données alimentaires.
 - Paramètres du nombre de population initiales : on a testé 4 valeurs distinctes : 3, 5, 7 et 9.
- Paramètre de génération aléatoire étudié et algorithme lstm : Taille de la recette : Nous avons varié la taille de la recette à générer en fonction de la taille minimale,

moyenne et maximale des recettes de notre ensemble de données réelles correspondant à 4,11 et 41, respectivement.

- Paramètre de l’algorithme LSTM : de plus de la taille de la recette, nous avons étudié l’impact de la taille de la séquence d’ingrédients qui alimente le lstm en entrée pour qu’il génère le reste de la recette. La taille de la séquence initiale a connu une variation de deux valeurs 3 et 4.

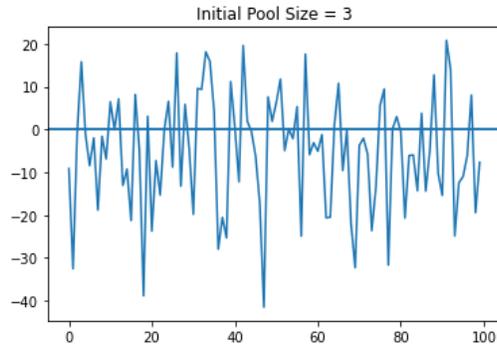
Avant d’effectuer cette étape, nous avons calculé le score de partage moyen des saveurs pour la cuisine maghrébine des recettes réelles présenté dans l’équation 2.3. Les valeurs résultantes étant $Ns(\text{real}) = 34.275600390776006$ pour l’ensemble de données de la cuisine maghrébine.

3.9.2 Résultats de l’algorithme génétique

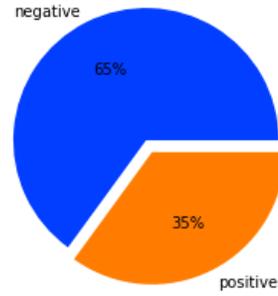
Tests et résultats sur la cuisine maghrébine Dans ce qui suit, nous allons tester plusieurs configurations de l’algorithme génétique (GA) avec diverses tailles de population initiale. Le tableau 3.6 résume les résultats des différentes configurations testées sur l’ensemble de données de la cuisine maghrébine. Les figures 3.6, 3.7, 3.8 et 3.9 montrent les variations de delta.

		Score	Observation
population initiale=3	min delta	-41.48	negative food
	max delta	20.84	pairing dans 65 %
	avrg delta	-5.15	des cas
population initiale=5	min delta	-55.06	négative food
	max delta	20.45	pairing dans 68 %
	avrg delta	-6.48	des cas
population initiale=7	min delta	-44.57	negative food
	max delta	21.70	pairing das 71 %
	avrg delta	-6.80	des cas
population initiale=9	min delta	-35.95	negative food
	max delta	19.81	pairing dans 64 %
	avrg delta	-6.15	des cas

TABLE 3.6 – Les résultats de GA avec différentes configurations sur les données de la cuisine maghrébine

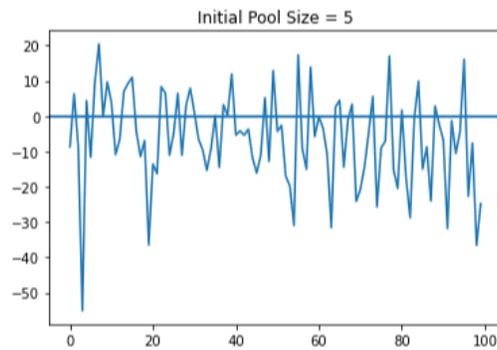


(a) Tracé de la variation du delta

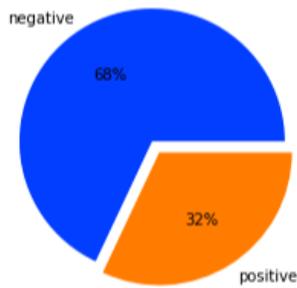


(b) Diagramme des résultats de food pairing

FIGURE 3.6 – Résultats de GA avec population initiale = 3, cuisine maghrébine

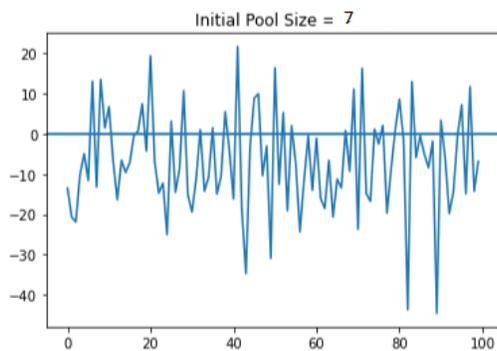


(a) Tracé de la variation du delta

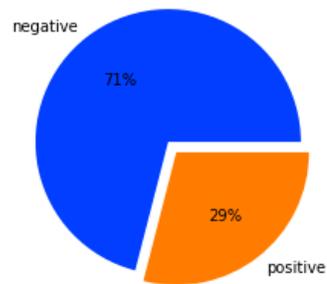


(b) Diagramme des résultats de food pairing

FIGURE 3.7 – Résultats de GA avec population initiale = 5, cuisine maghrébine

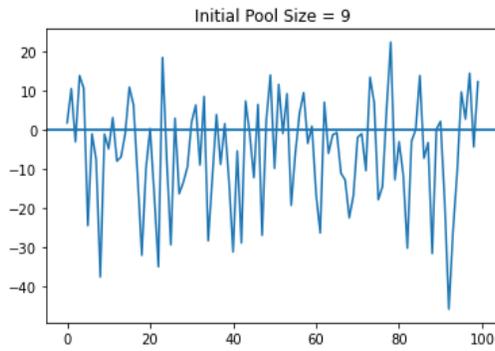


(a) Tracé de la variation du delta

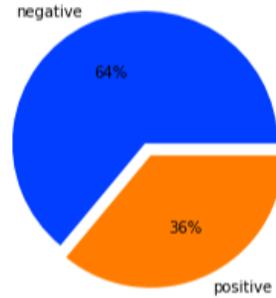


(b) Diagramme des résultats de food pairing

FIGURE 3.8 – Résultats de GA avec population initiale = 7, cuisine maghrébine



(a) Tracé de la variation du delta



(b) Diagramme des résultats de food pairing

FIGURE 3.9 – Résultats de GA avec une population initiale = 9, cuisine maghrébine

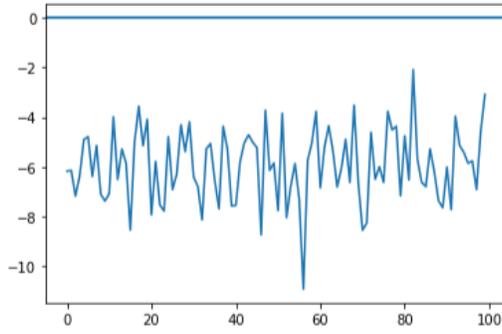
Toutes les configurations de GA (algorithme génétique) ont mené à la même conclusion en ce qui concerne les tendances de food pairing dans la cuisine maghrébine. Cette dernière présente une inclinaison négative avec un pourcentage moyen de 67.

3.9.3 Tests et résultats de la génération aléatoire

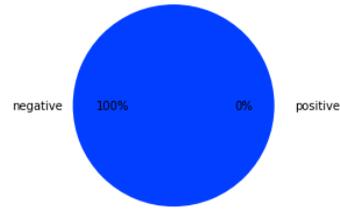
Ensemble de données de la cuisine maghrébine Nous avons générer les recettes synthétiques en variant la taille de la recette générer à 4, 11 et 41. Le tableau 3.7 résume les résultats obtenus en utilisant les différentes tailles de recettes sur l'ensemble de données de la cuisine maghrébine. Les figures 3.10, 3.11 et 3.12 montrent les variations de delta.

		Score	Observation
taille=4	min delta	-10.91	negative food pairing dans 100 % des cas
	max delta	-2.08	
	avrg delta	-5.93	
taille=11	min delta	-8.49	négative food pairing dans 100 % des cas
	max delta	-3.45	
	avrg delta	-5.99	
taille=41	min delta	-7.41	positive food pairing das 100 % des cas
	max delta	-4.78	
	avrg delta	-5.93	

TABLE 3.7 – Les résultats de génération aléatoire avec différentes configurations sur les données de la cuisine maghrébine

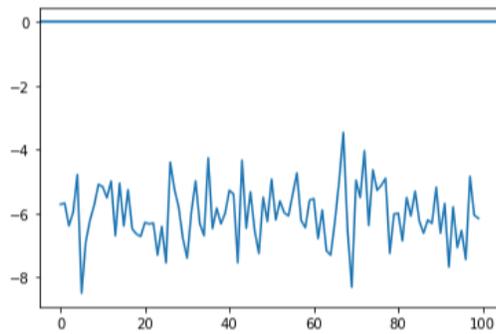


(a) Tracé de la variation du delta

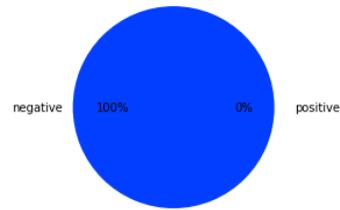


(b) Diagramme des résultats de food pairing

FIGURE 3.10 – Résultats de génération aléatoire avec une taille de recette = 4, cuisine maghrébine

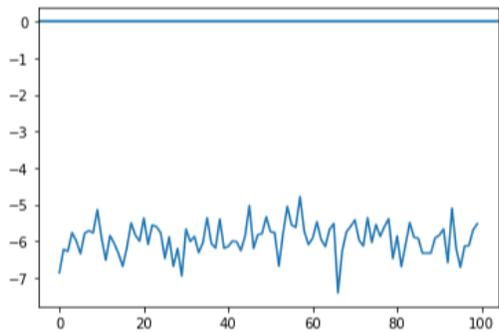


(a) Tracé de la variation du delta

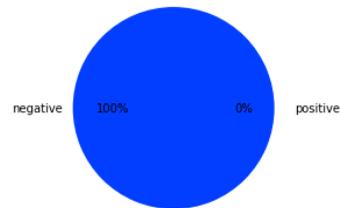


(b) Diagramme des résultats de food pairing

FIGURE 3.11 – Résultats de génération aléatoire avec une taille de recette = 11, cuisine maghrébine



(a) Tracé de la variation du delta



(b) Diagramme des résultats de food pairing

FIGURE 3.12 – Résultats de génération aléatoire avec une taille de recette = 41, cuisine maghrébine

Les trois différentes configurations de la génération aléatoires des recettes synthétiques

aboutissent à la même conclusion est que la cuisine maghrébine a une tendance de food pairing négative avec un pourcentage moyen de 100.

3.9.4 Tests et résultats de l'architecture LSTM

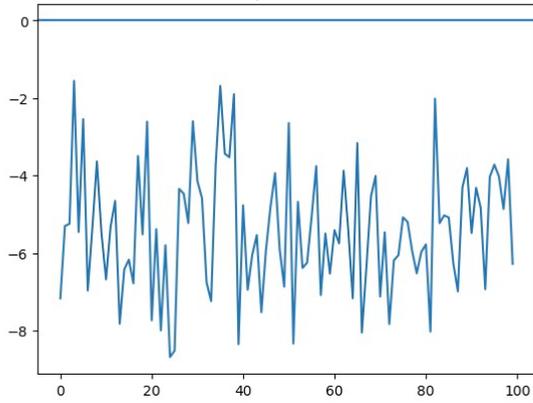
Pour l'architecture LSTM, nous avons utilisé générer de manière aléatoire la séquence initiale d'entrée de l'algorithme ayant pour taille N premiers ingrédients où $N \in \{3, 4\}$. Ensuite, nous avons laissé l'algorithme LSTM prédire le reste ingrédients pour compléter la taille de la recette. La taille de la recette a été variée à 4, 11 et 41, correspondant à la taille des recettes minimale, moyenne et maximale des recettes réelles constituant le corpus réel des recettes maghrébines.

Ensemble de données de la cuisine maghrébine, Paramètre de la taille la séquence initiale $N=3$

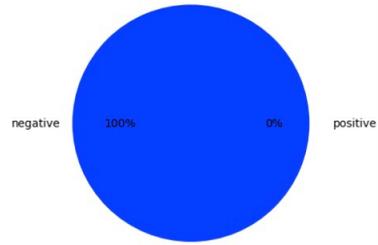
Les tests effectués dans cette section sont obtenus avec une taille de séquence initiale fixée 3. Cette séquence va alimenter le LSTM pour générer et compléter le reste de la recette synthétique. La taille de cette dernière a été variée et les résultats sont présentés dans le tableau 3.8. Les figures 3.13, 3.14 et 3.15 montrent les variations de delta.

		Score	Observation
taille=4	min delta	-8.68	negative food
	max delta	-1.55	pairing dans 100 %
	avrg delta	-5.40	des cas
taille=11	min delta	-11.26	négative food
	max delta	5.09	pairing dans 60 %
	avrg delta	-1.71	des cas
taille=41	min delta	-5.35	nègative food
	max delta	-2.77	pairing das 100 %
	avrg delta	-4.05	des cas

TABLE 3.8 – Les résultats de LSTM, taille recette 4, 11 et 41, cuisine maghrébine

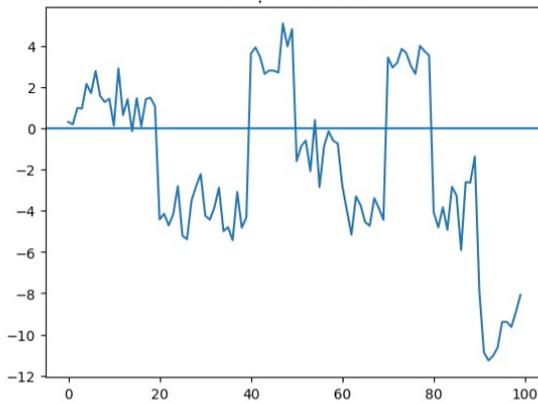


(a) Tracé de la variation du delta

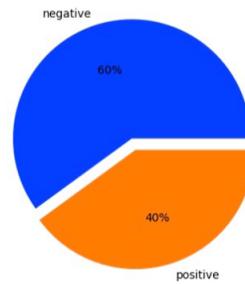


(b) Diagramme des résultats de food pairing

FIGURE 3.13 – Résultats de LSTM pour une taille de recette = 4, cuisine maghrébine

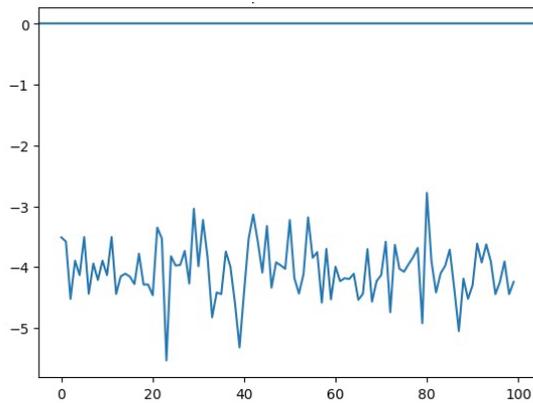


(a) Tracé de la variation du delta

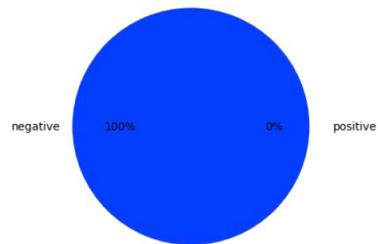


(b) Diagramme des résultats de food pairing

FIGURE 3.14 – Résultats de LSTM pour une taille de recette = 11, cuisine maghrébine



(a) Tracé de la variation du delta



(b) Diagramme des résultats de food pairing

FIGURE 3.15 – Résultats de LSTM pour une taille de recette = 41, cuisine maghrébine

Avec une taille de séquence initiale de $N=3$, le LSTM a généré des recettes qui ont permis de vérifier que la cuisine maghrébine a une tendance d'appariement d'ingrédients négative avec une moyenne de 80%.

Paramètre de la taille de la recette pour $N=4$

Dans cette section, la taille de la séquence d'entrée initiale qui alimente le LSTM sera fixée à 4. Le tableau 3.9 résume les résultats obtenus sur l'ensemble de données de la cuisine maghrébine en variant la taille des recettes synthétiques à générer par LSTM à 11 et 41. Les figures, 3.16 et ?? montrent les variations de delta.

		Score	Observation
taille=11	min delta	-7.18	négative food
	max delta	-3.03	pairing dans 100 %
	avrg delta	-5.09	des cas
taille=41	min delta	-9.02	nègarive food
	max delta	-7.11	pairing das 100 %
	avrg delta	-8.11	des cas

TABLE 3.9 – Les résultats de LSTM, taille de séquence $N=4$, cuisine maghrébine

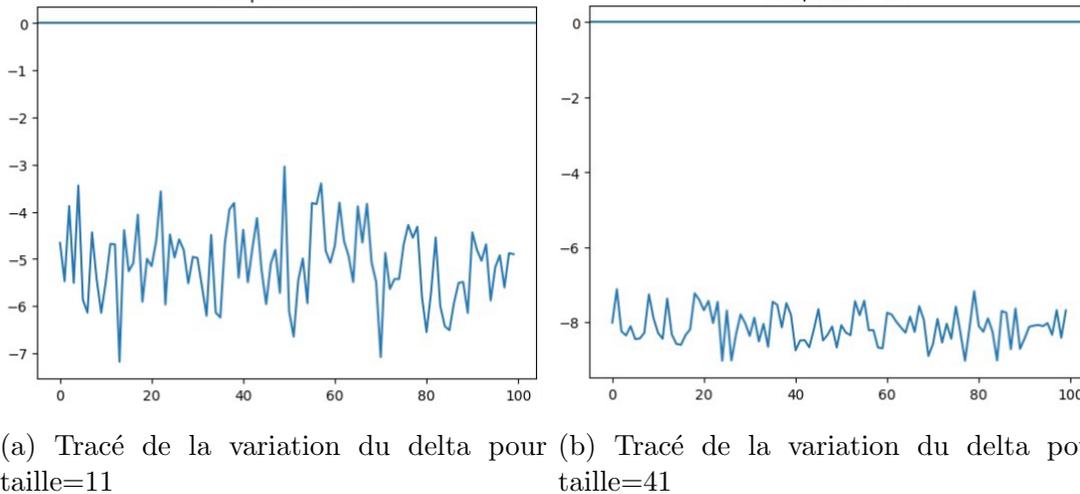


FIGURE 3.16 – Les résultats de LSTM, taille de séquence $N=4$, taille de recette = 11 et 41 cuisine maghrébine

Toutes les différentes configurations de l'algorithme LSTM testé ont mené à une tendance négative pour la cuisine maghrébine en ce qui concerne l'hypothèse du food pairing. Ces résultats convergent avec ceux obtenus par les autres méthodes de génération de recette synthétiques testés GA et génération aléatoire. La cuisine maghrébine a montré une inclinaison négative. Cette tendance est similaire à celle des pays du sud européen. Ceci indique que les peuples maghrébins et les peuples sud européens ont les mêmes goûts en ce qui concerne les choix culinaires.

3.10 Tests et résultats de l'architecture LSTM : étude de l'impact des épices

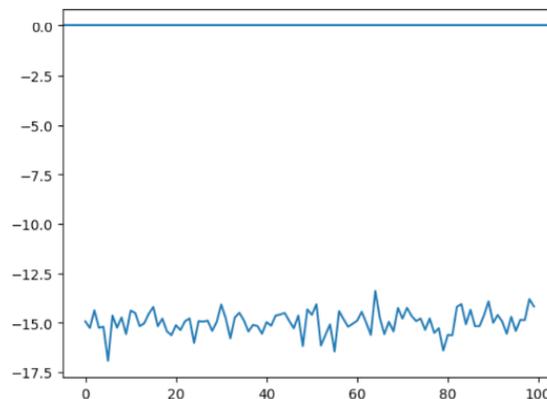
Afin d'étudier la contribution des épices à définir les tendances de food pairing dans la cuisine algérienne, nous allons générer des recettes aléatoires en fixants les premiers ingrédients par des épices et le reste des ingrédients de la recette est généré par le LSTM. Nous avons choisi deux compositions d'épices fréquemment utilisés dans la cuisine algérienne qui sont la **Dersa** et la **Chermoula**. Ces deux dernières sont la de plusieurs plats algériens tels que les chitithats et les chermoulats. De plus de ces compositions d'épices, nous avons fixé deux types de sauce qui sont à la base de la majorité des plats algériens : la sauce blanche et la sauce rouge. Cette étude nous permettra de déterminer si ces compositions d'épices et les types de sauces contribuent dans la tendance négative de la cuisine algérienne ou pas.

Dersa : [:'ail', 'cumin', 'piment rouge', 'sel', 'poivre rouge', 'poivre noir']

Le tableau 3.10 résume le résultat de configuration testée sur l'ensemble de données de la cuisine algérienne en fixant la composition des épices de la dersa et générant le reste des ingrédients des recettes en utilisant le LSTM. La figure 3.17, montre la variation de delta avec cette configuration de recettes.

		Score	Observation
taille=11	min delta	-16.91	négative food
	max delta	-13.39	pairing dans 100 %
	avrg delta	-14.97	des cas

TABLE 3.10 – Le résultat avec la composition dersa fixée sur la cuisine algérienne



(a) Tracé de la variation du delta

FIGURE 3.17 – Résultat pour une taille de recette = 11, composition dersa fixée dans les recettes générées

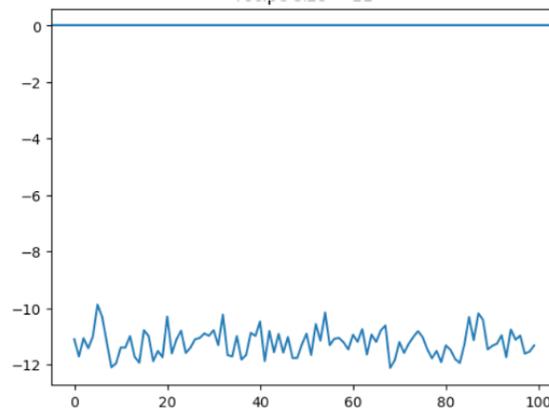
Le résultat obtenu montre que la composition des épices de la dersa algérienne contribue à 100% dans la tendance négative de food pairing de la cuisine algérienne. Ce résultat

suggère que les ingrédients sélectionnés initialement sont une des cause de l'inclinaison négative culinaire.

Chermoula : :['ail', 'coriandre', 'poivre rouge', 'poivre noir', 'citron', 'sel', 'huile'] Le tableau 3.11 résume le résultat de configuration testée sur l'ensemble de données de la cuisine algérienne en fixant les in grédients de la chermoula dans l'ensemble des recettes à générer par le LSTM. La figure 3.18, montre la variation de delta.

		Score	Observation
taille=11	min delta	-12.12	négative food
	max delta	-9.87	pairing dans 100 %
	avrg delta	-11.22	des cas

TABLE 3.11 – Le résultat de la fixation des ingrédients de la chermoula sur la cuisine algérienne



(a) Tracé de la variation du delta

FIGURE 3.18 – Résultat pour l'utilisation de la chermoula avec une taille de recette = 11

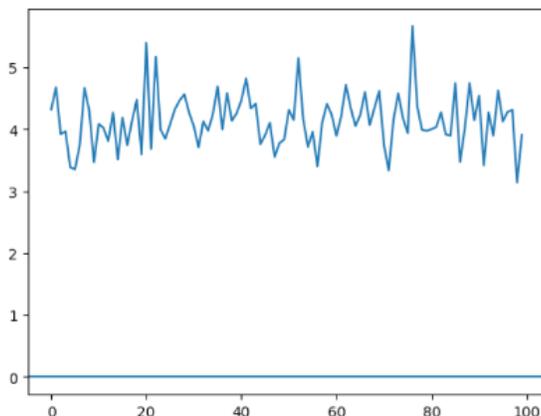
Le résultat obtenu confirme que les ingrédients de la chermoula contribue à 100% à la tendance négative de food pairing dans la cuisine algérienne.

Sauce rouge : ['tomate', 'smen', 'sel', 'poivre noir', 'poivre rouge', 'sucre', 'oignon']

Dans ce qui suit, nous allons générer dans recettes à base de sauce rouge algérienne et étudier son impact sur la tendance de food pairing. Le tableau 3.12 résume le résultat de configuration testée sur l'ensemble de données de la cuisine algérienne. La figure 3.19, montre la variation de delta.

		Score	Observation
taille=11	min delta	3.14	positive food
	max delta	5.66	pairing dans 100 %
	avrg delta	4.14	des cas

TABLE 3.12 – Le résultat de l’impact de la sauce rouge sur la cuisine algérienne



(a) Tracé de la variation du delta

FIGURE 3.19 – Résultat pour une taille de recette = 11, composition de la sauce rouge fixée

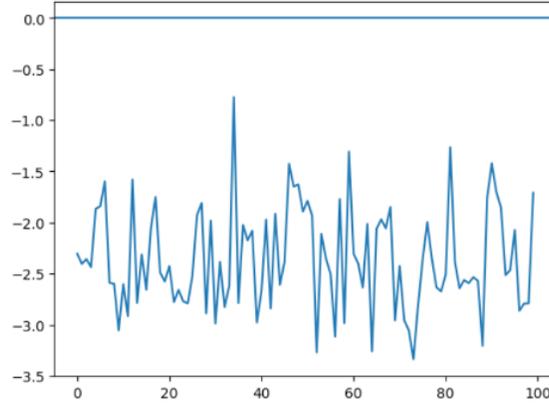
Les résultats présentés dans la figure 3.19 et 3.12 montrent que l’utilisation des recettes à base de sauce rouge algérienne mène vers une tendance positive. Ces résultats sont intéressants car ils montrent que la cuisine algérienne est très diverse. Tandis que les plats contenant à base de chermoula et dersa ont une tendance négative, ceux basés sur la sauce rouge ont une tendance positive.

Sauce blanche : [’oignon’, ’cannelle’, ’poivre noir’, ’sel’, ’huile’, ’persil’]

Sur la même lancée, nous effectuons des tests en générant des recettes à base de sauce blanche algérienne ou tester son impact sur le food pairing dans la cuisine algérienne. Le tableau 3.13 résume le résultat de configuration testée sur l’ensemble de données de la cuisine algérienne avec la composition de la sauce blanche. La figure 3.20, montre la variation de delta.

		Score	Observation
taille=11	min delta	-3.33	négative food
	max delta	-0.77	pairing dans 100 %
	avrg delta	-2.35	des cas

TABLE 3.13 – Le résultat de food pairing avec la sauce blanche algérienne



(a) Tracé de la variation du delta

FIGURE 3.20 – Résultat de l’utilisation de la sauce blanche avec une taille de recette = 11

Le résultat obtenu confirme que la sauce blanche contribue dans la tendance négative de food pairing dans la cuisine algérienne à 100%.

3.11 L’étude du food pairing dans la cuisine algérienne

Un travail réalisé en 2022 par Kerbedj et al.[27] portait sur l’étude l’hypothèse de food pairing dans la cuisine algérienne en se basant sur les algorithmes génétiques. Les résultats obtenus par ce travail ont montré que la cuisine algérienne a une tendance négative. Afin d’enrichir ce travail et étudier la convergence des résultats de l’année précédente et de cette année concernant la cuisine algérienne, nous allons effectuer des test en utilisant différents algorithmes pour le processus de génération : algorithme génétique proposé dans le travail [27], une génération aléatoire et l’algorithme LSTM que nous avons proposé. Ces algorithmes seront appliquées sur des données algériennes provenant de différentes sources :

- Ensemble 1 : Ensemble de données extraits à partir du livre de Salima Hadjiat [29] utilisé dans ce travail. Ce livre contient des recettes de la région maghrébine. Nous allons extraire uniquement les recettes qui proviennent de l’Algérie afin de construire un premier sous ensemble de test.
- Ensemble 2 : Les recettes extraites du livre de Mme Hadjiat [29] seront combinées avec l’ensemble de recettes collectées, à partir du livre de Mme Bouayed, utilisées dans le travail [27] pour construire un corpus de recettes algériennes authentiques enrichies.

3.11.1 Résultats de l’hypothèse du food pairing sur ensemble 1 de donnée algérienne

Nous allons évaluer l’ensemble 1 des recettes algériennes extraites exclusivement à partir du livre Mme Hadjiat [29] avec les trois méthodes de génération de recettes : les algorithmes

génétiques, une génération aléatoire et l’algorithme LSTM et voir si les résultats sur ce nouveau ensemble de données convergent avec les résultats obtenus dans le travail [27] où la cuisine algérienne avait une tendance négative.

Le tableau 3.14 résume les résultats obtenus sur l’ensemble 1 en utilisant les algorithmes génétiques. Il montre différentes configurations testées sur l’ensemble 1 des données de la cuisine algérienne. Ces résultats montrent que la cuisine algérienne a maintenu la tendance négative sur ce nouveau ensemble de données avec une inclinaison négative moyenne de 83%.

		Score	Observation
population initiale=3	min delta	-47.15	négative food
	max delta	11.46	pairing dans 79 %
	avrg delta	-13.19	des cas
population initiale=5	min delta	-47.73	négative food
	max delta	14.63	pairing dans 79 %
	avrg delta	-10.42	des cas
population initiale=7	min delta	-41.72	negative food
	max delta	9.95	pairing dans 88 %
	avrg delta	-14.63	des cas
population initiale=9	min delta	-53.99	negative food
	max delta	11.05	pairing dans 87 %
	avrg delta	-14.16	des cas

TABLE 3.14 – Les résultats des algorithmes génétiques sur l’ensemble de données algériennes extraites du livre de Mme Hadjiat

Par la suite et pour le même ensemble de données, nous avons changé les algorithmes génétiques par une génération aléatoire des recettes avec différentes tailles de recette et avons comparé les résultats qui sont illustrés dans le tableau 3.15.

		Score	Observation
taille=4	min delta	-17.67	negative food
	max delta	-9.54	pairing dans 100 %
	avrg delta	-14.18	des cas
taille=11	min delta	-15.90	négative food
	max delta	-11.22	pairing dans 100 %
	avrg delta	-13.81	des cas
taille=41	min delta	-15.01	positive food
	max delta	-12.89	pairing dans 100 %
	avrg delta	-13.97	des cas

TABLE 3.15 – Les résultats obtenus par la génération aléatoire sur l’ensemble de recettes algériennes de Mme Hadjiat

Les résultats par la génération aléatoire obtenus sur l’ensemble 1 convergent avec ceux obtenus par les algorithmes génétiques en ce qui concerne la tendance négative de food

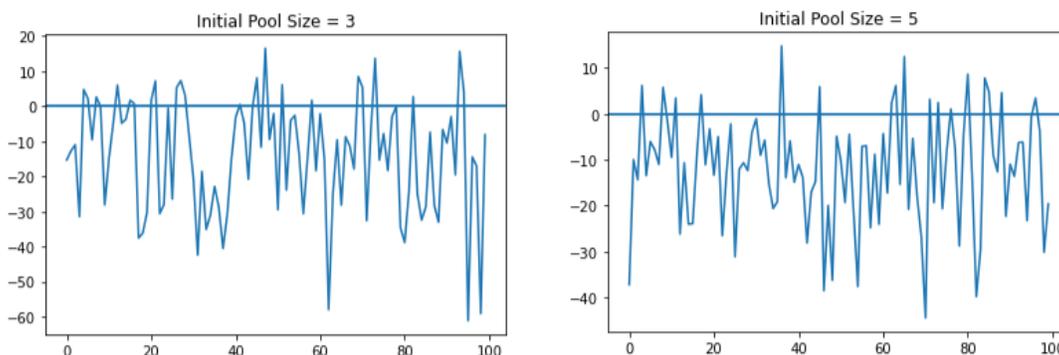
pairing dans la cuisine algérienne. Il reste de tester l’algorithme LSTM sur le même ensemble de données pour confirmer les deux conclusion précédentes. Le tableau 3.16 résume les résultats des différentes configurations testées avec l’algorithme LSTM.

		Score	Observation
taille=4	min delta	-17.06	negative food
	max delta	-9.52	pairing dans 100 %
	avrg delta	-13.24	des cas
taille=11	min delta	-16.17	négative food
	max delta	-10.15	pairing dans 100 %
	avrg delta	-13.30	des cas
taille=41	min delta	-14.33	nègative food
	max delta	-12.01	pairing das 100 %
	avrg delta	-13.17	des cas

TABLE 3.16 – Les résultats de l’algorithme LSTM sur l’ensemble de données algériennes extraites du livre de Mme Hadjiat

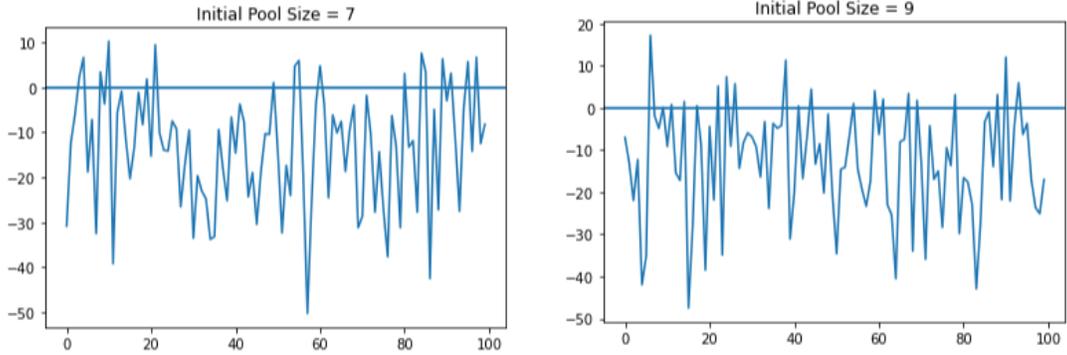
3.11.2 Résultats de l’hypothèse du food pairing sur ensemble 2 de donnée algérienne

Dans cette partie, nous allons construire un ensemble de recettes algériennes authentiques enrichie en combinant les recettes extraites à partir du livre de Mme Bouayed et celles du livre de Mme Hadjiat. Sur ce nouveau ensemble de recettes enrichi et afin d’étudier en profondeur l’hypothèse du food pairing dans la cuisine algérienne, nous allons tester trois processus de génération à base de l’algorithme génétique, génération aléatoire et l’algorithme LSTM. Les résultats des expérimentations sont illustrés dans les figures , Les figures 3.21 et 3.22 montrent les variations de delta, La figure 3.23, 3.24 montre la variation de delta, respectivement.



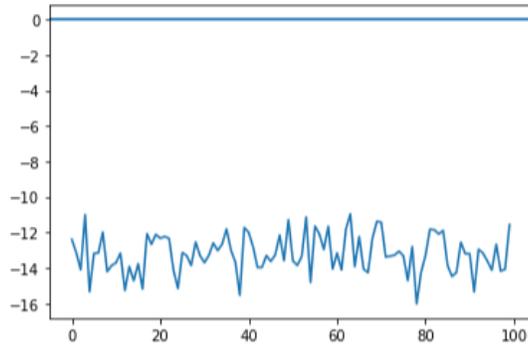
(a) Tracé de la variation du delta pour pop=3 (b) Tracé de la variation du delta pour pop=5

FIGURE 3.21 – Résultats des GA sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une population initiale = 3 et 5



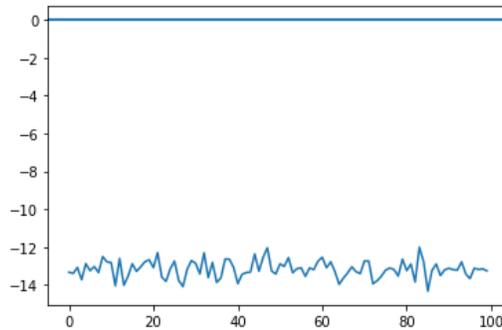
(a) Tracé de la variation du delta pour pop=7 (b) Tracé de la variation du delta pour pop=9

FIGURE 3.22 – Résultats des GA sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une population initiale = 7 et 9



(a) Tracé de la variation du delta

FIGURE 3.23 – Résultat de la génération aléatoire sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une taille de recette = 11



(a) Tracé de la variation du delta

FIGURE 3.24 – Résultat de l’algorithme LSTM sur les données extraites des livres de Mme Hadjiat et Mme Bouayed, pour une taille de recette = 11

Les résultats obtenus sur l'ensemble 2 qui combine les recettes algériennes des livres [29] et [39] avec les trois méthodes de génération de recettes : algorithme génétique, génération aléatoire et l'algorithme LSTM confirment que la cuisine algérienne a une tendance négative élevée avec une moyenne $\approx 80\%$ concernant le choix de l'appariement des ingrédients. Cela indique que les ingrédients sélectionnés dans un même plat algérien partagent rarement des molécules commun. Cette analyse approfondie nous permet de constater que les résultats sur les données algériennes s'alignent parfaitement avec les résultats obtenus dans le travail de Kerbedj [27] sur la tendance de la cuisine algérienne.

3.12 Conclusion

Dans ce chapitre, nous avons décrit toutes les étapes suivies pour mettre en pratique notre étude de l'hypothèse de food pairing dans la cuisine maghrébine. En commençant par la méthode de collecte du corpus de recettes maghrébines et les techniques de prétraitement que nous avons employées, tous les tests que nous avons effectués pour vérifier l'exactitude de notre conclusion, et enfin en parlant des différentes configurations des approches suggérées et leurs résultats. Par ailleurs, nous avons étudié l'impact et la contribution des épices dans la tendance négative de la cuisine algérienne et nous avons comparé nos résultats à ceux obtenus dans le travail [27] portant sur le food pairing dans la cuisine algérienne. Pour effectuer cette comparaison, nous avons créé deux nouveaux sous ensembles de recettes algériennes authentiques. Les résultats obtenus ont appuyé la conclusion sur la tendance négative de la cuisine algérienne et qui au final converge avec la cuisine maghrébine.

Conclusion générale

En définissant les caractéristiques du food pairing qui le distinguent selon l'hypothèse de Blumenthal, et en explorant ce que cela révèle sur les origines et les influences de la cuisine maghrébine, cette thèse a cherché à intégrer le domaine du food pairing dans la gastronomie maghrébine.

Nos recherches ont montré que la cuisine maghrébine présente une tendance négative en matière de food pairing, c'est-à-dire que les aliments qui sont couramment combinés dans les plats ne partagent pas beaucoup de molécules aromatiques. Nous avons également remarqué que la tendance de la cuisine du sud de l'Europe (Méditerranéenne européenne) est défavorable, ainsi que la cuisine algérienne selon l'étude qui a été faite l'année passée, ce qui suggère que les traditions culinaires et les goûts de ces trois régions adjacentes sont comparables.

Nous avons rassemblé des recettes maghrébines à partir d'un livre écrit par Salima HADJIAT, puis nous les avons utilisées pour parvenir à cette conclusion. En utilisant des techniques de traitement du langage naturel, nous avons identifié les composants distinctifs de chaque recette et les avons mis en corrélation avec les autres recettes.

Les recettes maghrébines ont été soumises à un profil de saveur incluant les composés aromatiques qui leur confèrent leur saveur et leur goût. À partir des saveurs, nous pouvons examiner les éléments chimiques que les aliments ont en commun. L'examen de l'espace des composants créés nécessite l'utilisation de l'architecture LSTM pour optimiser la recherche.

Notre conclusion résout le principe fondamental de food pairing et ouvre la voie à des points de vue créatifs sur la cuisine maghrébine. La solution à cette question débouchera sur un nouveau partenariat entre scientifiques et cuisiniers maghrébins qui lancera l'univers des relations improbables.

Les informaticiens peuvent intervenir pour appliquer des techniques d'apprentissage profond afin de rationaliser le processus lorsqu'ils sont confrontés à un grand nombre d'ingrédients et à une infinité de combinaisons à explorer à la recherche des paires idéales, dans notre cas des ingrédients ayant le moins de composants chimiques en commun. Les chefs peuvent alors tirer parti de ces connaissances pour élaborer un plat avec le dosage et l'équilibre appropriés. Surtout, il ne faut pas oublier que la cuisine est un art, une expression humaine sensible, et que la logique scientifique ne doit pas obliger la gastronomie à s'en écarter, sur le plan gustatif et sensoriel.

À partir de notre étude, nous avons constaté une tendance négative dans la cuisine maghrébine, ce qui ouvre la voie à une nouvelle alliance entre chefs maghrébins et scienti-

fiques pour explorer l'univers des associations improbables. Les résultats positifs de l'utilisation de la sauce blanche dans la cuisine algérienne soulignent l'importance d'étudier en profondeur les variations des tendances culinaires entre les différentes régions du Maghreb. Dans cette perspective, les informaticiens peuvent jouer un rôle clé en utilisant des techniques d'apprentissage profond pour optimiser la recherche de nouvelles combinaisons innovantes d'ingrédients. En se concentrant sur les ingrédients ayant le moins de composants chimiques en commun, ils peuvent identifier des paires d'ingrédients offrant des associations de saveurs contrastées et uniques. Cette collaboration entre chefs maghrébins et scientifiques ouvrirait de nouvelles possibilités dans la gastronomie maghrébine, en créant des alliances culinaires inattendues qui pourraient surprendre et ravir les papilles des amateurs de cuisine.

Références

- [1] Sanatan Mishra. Unsupervised learning and data clustering, 2017. Site.
- [2] James P.Bagrow et al. Yong-Yeol Ahn, Sebastian E.Ahnert. Flavor network and the principles of food pairing. 2011. Article.
- [3] Taha Alfakih Muna Al-Razgan, Shahad Tallab. Exploring the food pairing hypothesis in saudi cuisine using genetic algorithm. 2021. Article.
- [4] About the industry and food service careers. Site.
- [5] Merriam. Merriam-webster, Official Website on www.Merriam-Webster.com. Site, last accessed August 2022.
- [6] Sebastian E Ahnert. Network analysis and data mining in food science : the emergence of computational gastronomy. 2013. Article.
- [7] Muna Saleh Alrazgan Shahad Talal Tallab. Exploring the food pairing hypothesis in arab cuisine : A study in computational gastronomy. 2016. Article.
- [8] Flavordb, Official Website on www.FlavorDB.com. Site.
- [9] Rudraksh Tuwani et al Neelansh Garg, Apuroop Sethupathy. Flavordb : a database of flavor molecules. 2017. Article.
- [10] J. Wang K. R. Varshney, L. R. Varshney and D. J. Myers. Flavor pairing in medieval european cuisine : a study in cooking with dirty data, 2013. Article.
- [11] A. Jain and G. J. Bagler. Spices form the basis of food pairing in indian cuisine, 2015. Article.
- [12] Heston Blumenthal. Weird but wonderful. 2002. Article.
- [13] Xavier Giro-i-Nieto Adriana Romero Amaia Salvador, Michal Drozdal. Inverse cooking : Recipe generation from food images. 2019. Article.
- [14] Jun Wang Yong Yu† Lantao Yu, Weinan Zhang†. Seqgan : Sequence generative adversarial nets with policy gradient. 2017. Article.
- [15] Xuhong Hou Bin Sheng Siyuan Pan, Ling Dai. Chefgan : Food image generation from recipes, 2020. Article.
- [16] Avijit Verma Bappaditya Mandal, N. B. Puan. Deep convolutional generative adversarial network based food recognition using partially labeled data, 2018. Article.
- [17] Palakorn Achananuparp et al Helena H. Lee, Ke Shu. Recipegpt : Generative pre-training based cooking recipe generation and evaluation system, 2020. Article.

- [18] Kai Fan et al Yizhe Zhang, Zhe Gan. Adversarial feature matching for text generation. 2017. Article.
- [19] Laurynas Karpus et al Donatas Repecka, Vykintas Jauniskis. Expanding functional protein sequence space using generative adversarial networks. 2019. Article.
- [20] Andrew DeLong et al Nathan Killoran, Leo J. Lee. Generating and designing dna with deep generative models. 2017. Article.
- [21] Eoin Brophy John Weldon, Tomas Ward. Generation of synthetic electronic health records using a federated gan. 2021. Article.
- [22] Artur d’Avila Garcez Charitos Charitou, Simo Dragicevic. Synthetic data generation for fraud detection using gans. 2021. Article.
- [23] Theophile Gaudin et al Philippe Schwaller, Teodoro Laino. Molecular transformer : A model for uncertainty-calibrated chemical reaction prediction. 2019. Article.
- [24] Kuzma Khrabrov et al Artur Kadurin, Sergey Nikolenko. drugan : An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. 2017. Article.
- [25] Vu Thi Ngoc Anh et al Dong Do Duc, Phuc Thai Dinh. An efficient ant colony optimization algorithm for protein structure prediction. 2018. Article.
- [26] Shuib Basri Ammar K Alazzawi, Helmi Md Rais. Abcvs : An artificial bee colony for generating variable t-way test sets. 2019. Article.
- [27] CHAHBOUB Racha KERBEDJ Tarek. Etude de l’hypothèse de food pairing deans la cuisine algerienne. 2022. Article.
- [28] Raphaël Richard. Long short term memory, 2019. Site, dernier accès Avril 2023.
- [29] Salima Hadjiat. *La cuisine maghrèbine*. 1987. Book.
- [30] Essentials of deep learning : Introduction to long short term memory, December 10, 2017. Site.
- [31] Python introduction. Site.
- [32] re—regular expression operations—python 3.10.5 documentation. Site, dernier accès juin 2023.
- [33] pandas - python data analysis library. Site , dernier accès juin 2023.
- [34] Natural language toolkit. Site , dernier accès juin 2023.
- [35] Plotly python graphing library. Site , dernier accès juin 2023.
- [36] Reena Gupta Akshita Mehta. Egg yolk. 2021. Article.
- [37] Biology of eggs. 2021. Article.
- [38] PhytothÉrapie : Anis vert. 2012. Article.
- [39] Bouayed Fatima Zohra. *La cuisine algerienne*. Janvier 1970. Book.