



EXERSYS : Un système de recommandation appliqué à la nutrition, combinant graphes de connaissances et apprentissage automatique

Soutenance finale
5 septembre 2023

Présenté par: Noémie Jacquet
Dans le cadre de: Stage de 6 mois – Certificat de Spécialité IODAA

Maître de stage: Mme Cristina MANFREDOTTI
Enseignant tuteur: Mme Liliana IBANESCU

Contexte et problématique du stage

Au sein de l'UMR **MIA Paris-Saclay**



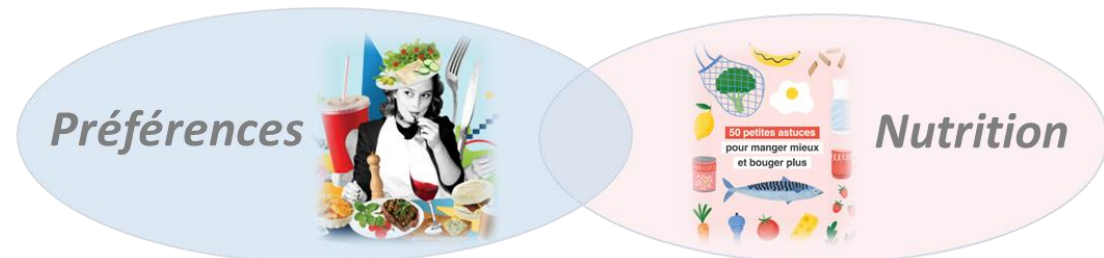
Sur le projet **Exersys** :

➤ Interdisciplinaire

Experts

- **Mme Cristina Manfredotti**, Maître de conférences -Ekinocs -AgroParisTech -INRAE -Univ. Paris Saclay.
- **M. Vincent Guigue**, HDR, Professeur -Ekinocs -AgroParisTech -INRAE -Univ. Paris Saclay.
- **M. Stéphane Dervaux**, Ingénieur en informatique -Ekinocs -AgroParisTech - INRAE – Univ. Paris Saclay
- **Mme Fatiha Saïs**, HDR, Professeur - LaHDAK - LISN -Univ. Paris Saclay.
- **M. Paolo Viappiani**, Chercheur – CNRS – LAMSADE -Univ. Paris Dauphine.
- **M. Nicolas Darcel**, HDR, Maître de conférences - PNCA - AgroParisTech - INRAE -Univ. Paris Saclay

➤ Visant à développer un **système de recommandation** alimentaire qui satisfait les préférences des utilisateurs et les contraintes nutritionnelles

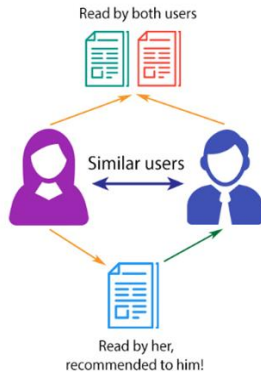


Objectif d'EXERSYS

élaborer un système de recommandation de repas combinant:

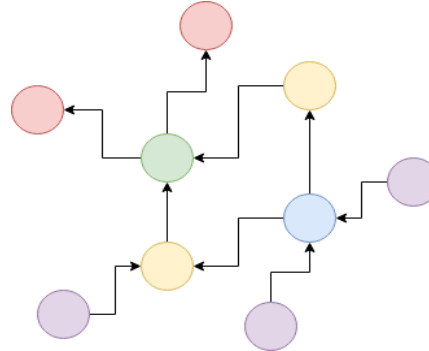
1 Axe 1:

Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**



2 Axe 2:

Intégration de **contraintes nutritionnelles** grâce aux **graphes de connaissance**



Etat de l'art:

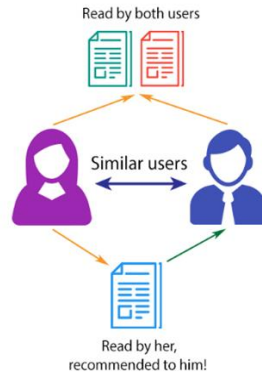
combinaison des 2 méthodes: **peu** de publications, **aucune alimentaire** dans le domaine

Source des données

1

Axe 1:

Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

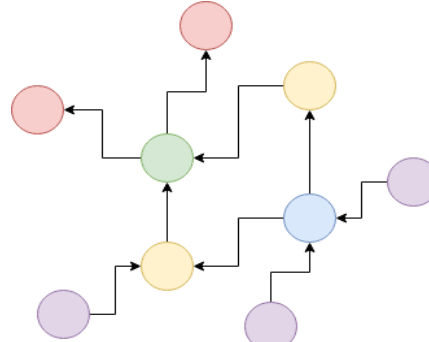


enquête alimentaire
nationale (7j)

2

Axe 2:

Intégration de **contraintes nutritionnelles** grâce aux **graphes de connaissance**



Connaissance
Expert



Périmètre de mon travail au sein du projet EXERSYS

Axe 1: 2 modèles testés

1 Axe 1:

Recommandation basée sur les
préférences utilisateurs par
apprentissage automatique

- ❑ Approche type filtrage collaboratif: '**FiltreCollab**'
- ❑ Génération de séquences par Réseau de neurones récurrents: '**GenSeqRNN**'

Périmètre de mon travail au sein du projet EXERSYS

Axe 1: 2 modèles testés

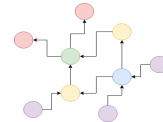
1 Axe 1:

Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

- ❑ Approche type filtrage collaboratif: '**FiltreCollab**'
- ❑ Génération de séquences par Réseau de neurones récurrents: '**GenSeqRNN**'

2 Axe 2:

Intégration de **contraintes nutritionnelles** grâce aux **graphes de connaissance**



Axe 2 = travaux d'Ayoub Hammal (stagiaire)

Périmètre de mon travail au sein du projet EXERSYS

Axe 1: 2 modèles testés

Connexion avec l'axe 2
testée sur un modèle

1 Axe 1:

Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

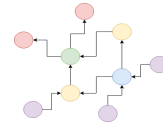
- ❑ Approche type filtrage collaboratif: '**FiltreCollab**'
- ❑ Génération de séquences par Réseau de neurones récurrents: '**GenSeqRNN**'



Génération de recommandation

2 Axe 2:

Intégration de **contraintes nutritionnelles** grâce aux **graphes de connaissance**



Axe 2 = travaux d'Ayoub Hammal (stagiaire)

Axe 1: 2 modèles

1 Axe 1:

Recommandation basée sur les
préférences utilisateurs par
apprentissage automatique

- ❑ Approche type filtrage collaboratif: **'FiltreCollab'**



- ❑ Génération de séquences par Réseau de neurones récurrents: **'GenSeqRNN'**



Axe 1: le modèle '*FiltreCollab*'

1

Axe 1:

Recommandation basée sur les
préférences utilisateurs par
apprentissage automatique

- ❑ Approche type filtrage collaboratif: '*FiltreCollab*'



- ❑ Génération de séquences par Réseau de neurones récurrents: '*GenSeqRNN*'

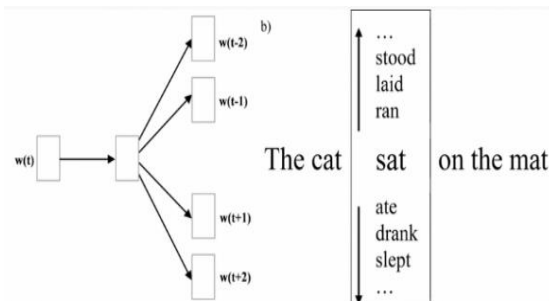
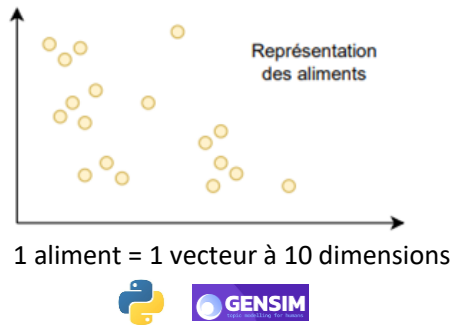


Modèle '*FiltreCollab*' développé sur le petit déjeuner: Méthodologie

PDJ

1. Apprendre un espace de représentation des aliments

Word2Vec entraîné sur les petits déjeuners des utilisateurs

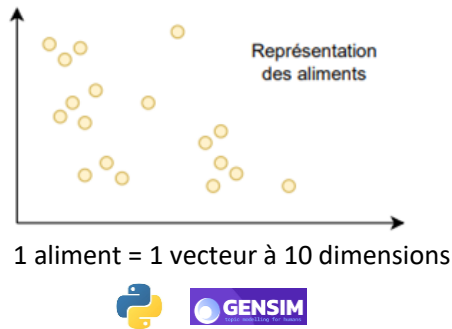


Modèle '*FiltreCollab*' développé sur le petit déjeuner: Méthodologie



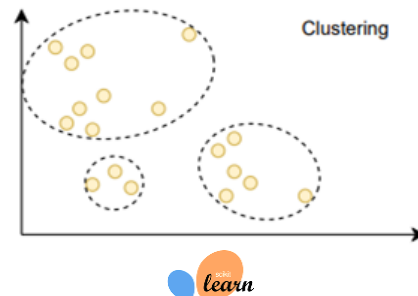
1. Apprendre un espace de représentation des aliments

Word2Vec entraîné sur les petits déjeuners des utilisateurs



2. Catégories d'aliments proches

K-means, k catégories

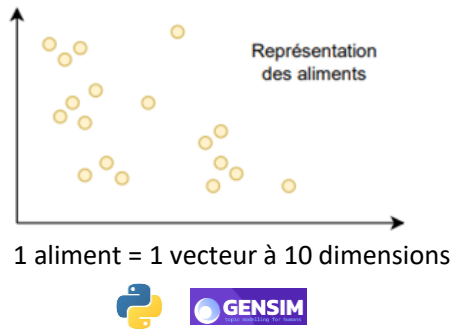


Modèle '*FiltreCollab*' développé sur le petit déjeuner: Méthodologie



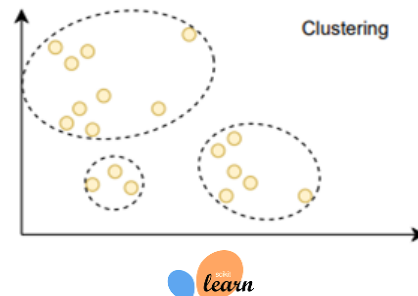
1. Apprendre un espace de représentation des aliments

Word2Vec entraîné sur les petits déjeuners des utilisateurs



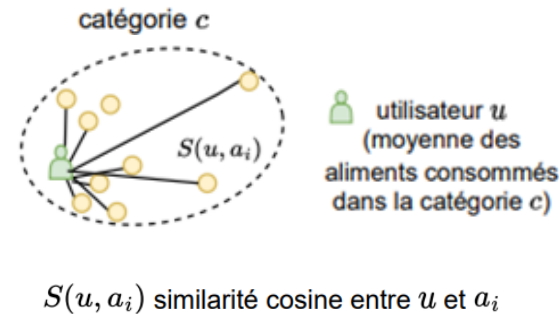
2. Catégories d'aliments proches

K-means, k catégories



3. Modélisation et tirage des aliments par catégorie

Recommandation d'un aliment : Tirage selon loi multinomiale



$$P(a_1 | c) = \frac{e^{\alpha \cdot S(u, a_1)}}{\sum_i e^{\alpha \cdot S(u, a_i)}}$$

c catégorie consommée

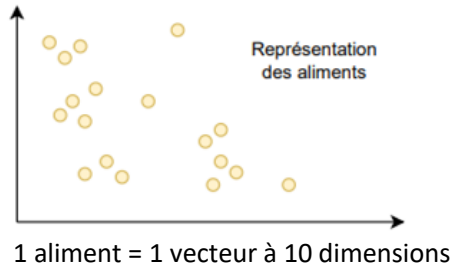
Modèle '*FiltreCollab*' développé sur le petit déjeuner: Méthodologie

Génération de recommandation

PDJ

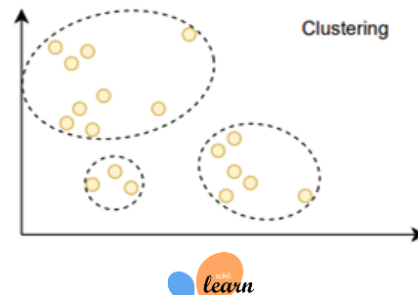
1. Apprendre un espace de représentation des aliments

Word2Vec entraîné sur les petits déjeuners des utilisateurs



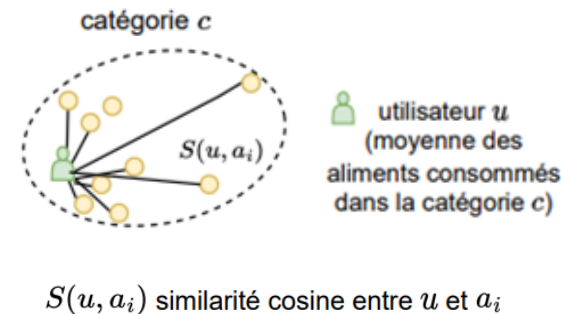
2. Catégories d'aliments proches

K-means, k catégories



3. Modélisation et tirage des aliments par catégorie

Recommandation d'un aliment : Tirage selon loi multinomiale



$$P(a_1 | c) = \frac{e^{\alpha \cdot S(u, a_1)}}{\sum_i e^{\alpha \cdot S(u, a_i)}}$$

c catégorie consommée

4. Recommandation d'un repas

Recommandation d'un repas:

- **Modalités de tirage**
(1 aliment par catégorie déjà consommée et dans $\beta\%$ des cas sinon)
- **Règles d'association/exclusion** d'aliments

Modèle *'FiltreCollab'* : critères d'évaluation appliqués



Génération de recommandation

1. Apprendre un espace de représentation des aliments

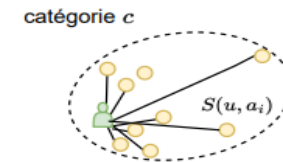
Word2Vec entraîné sur les repas

2. Catégories d'aliments proches

Partitionnement

3. Modélisation et tirage des aliments par catégorie

Recommandation d'un aliment :
Tirage selon loi multinomiale
(probabilités fonction des distances)



4. Recommandation d'un repas

Recommandation d'un repas:
➤ Modalités de tirage spécifiques
➤ Et règles d'association/exclusion d'aliments

Critères d'évaluation

Pertinence de l'espace appris

Interchangeabilité des aliments au sein d'une catégorie

Pertinence du positionnement des aliments au sein des catégories:

→ Log-vraisemblance
→ MRR (Rang moyen réciproque)

Cohérence d'un repas

(à quantifier dans 1 étude auprès d'utilisateurs)

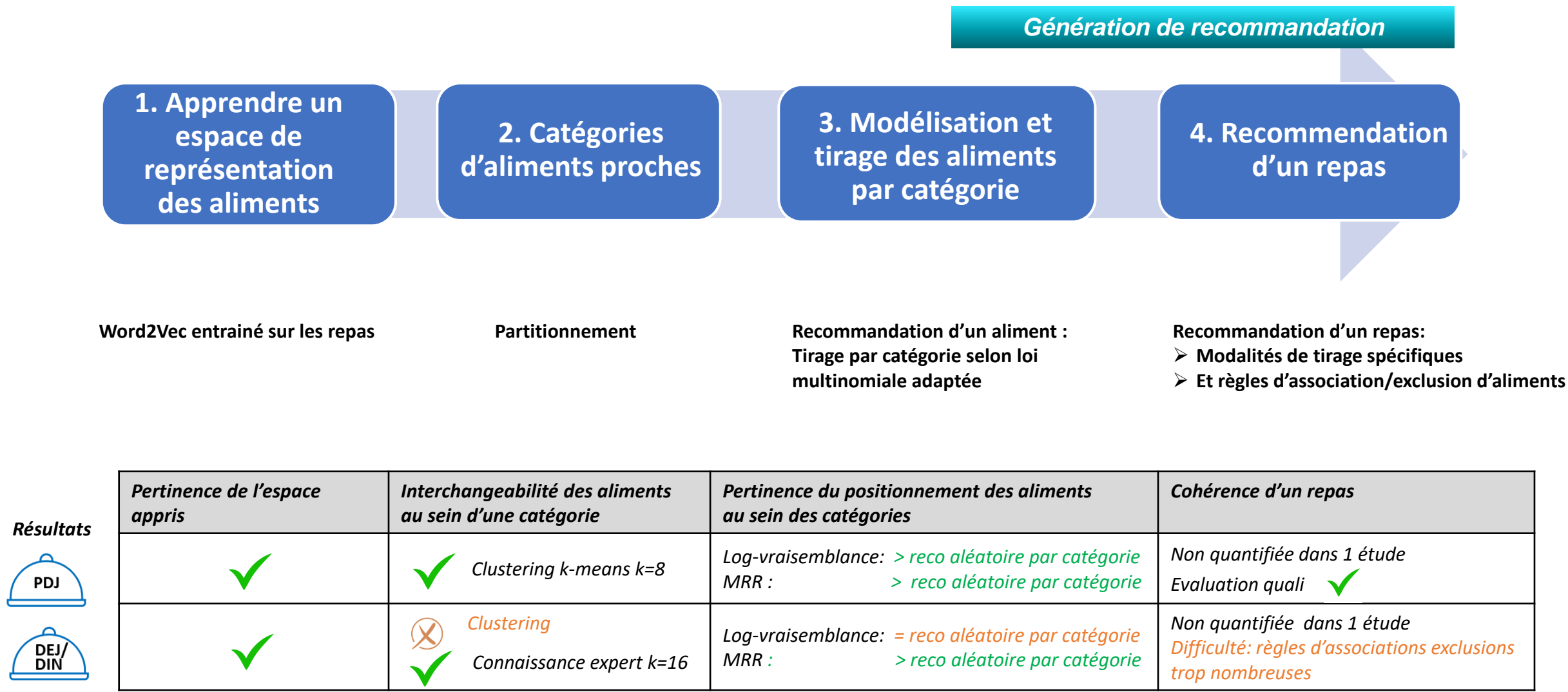
Ensemble d'apprentissage
J1 à J6



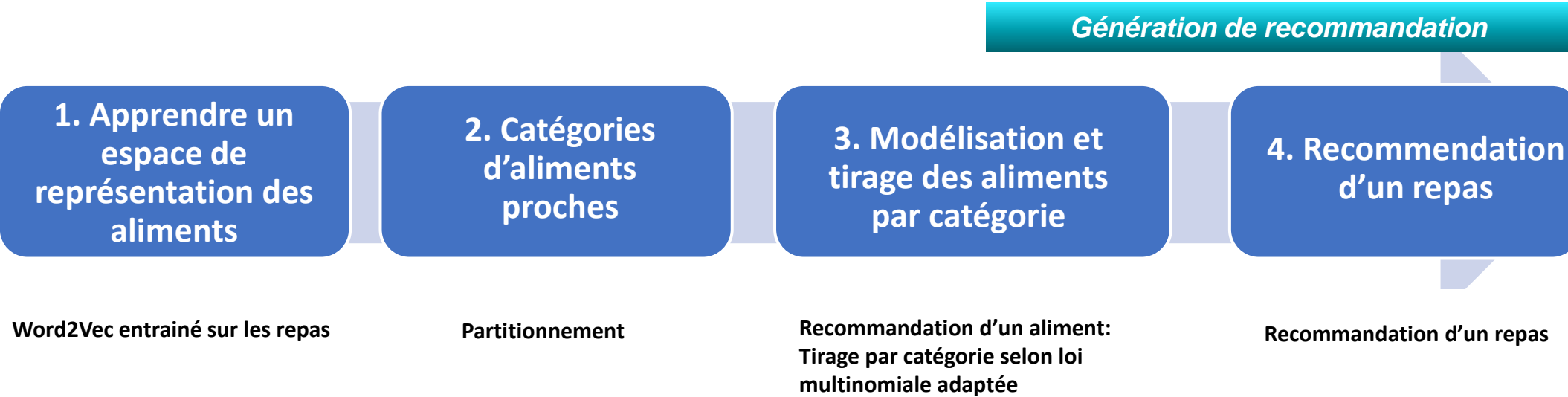
Ensemble de test
J7 jour masqué



Modèle *'FiltreCollab'*: performance sur le déjeuner/diner vs le petit déjeuner



Modèle *'FiltreCollab'*: moindre performance sur le déjeuner/diner



Résultats

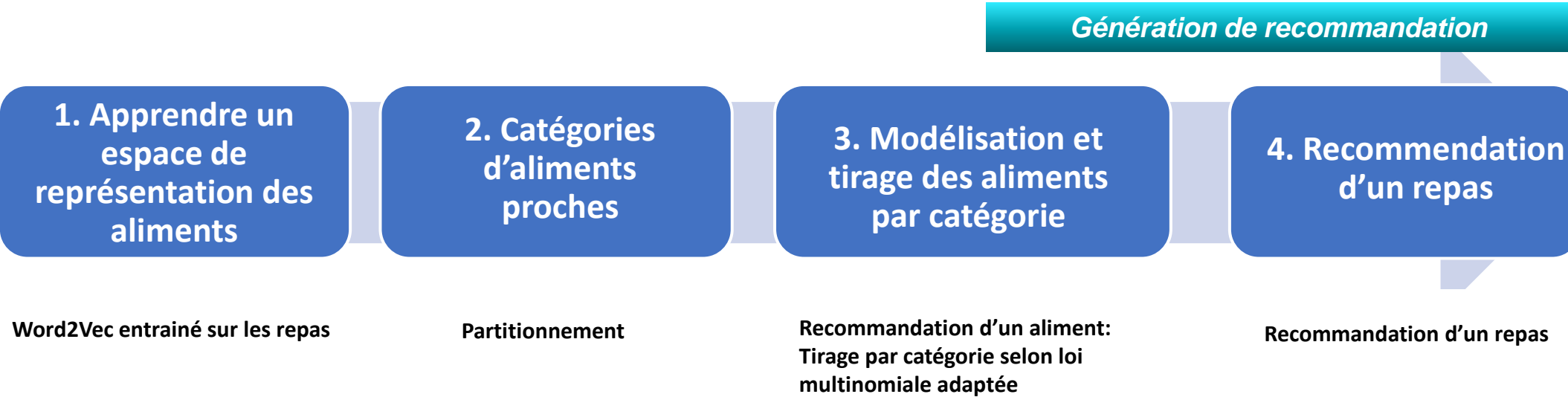


Interchangeabilité des aliments au sein d'une catégorie	Pertinence du positionnement des aliments au sein des catégories	Cohérence d'un repas
<div></div> <div>Clustering Connaissance expert k=16</div>	<div>Log-vraisemblance: = reco aléatoire par catégorie MRR : > reco aléatoire par catégorie</div>	<div>Non quantifiée dans une étude Difficulté: règles d'associations exclusions trop nombreuses</div>

Performance moindre:



- *Positionnement des aliments (ordre et non distances absolues)*
- *On bénéficie moins de l'historique*

Modèle *'FiltreCollab'*: moindre performance sur le déjeuner/diner



Résultats



Interchangeabilité des aliments au sein d'une catégorie	Pertinence du positionnement des aliments au sein des catégories	Cohérence d'un repas
<div> <i>Clustering</i></div> <div> Connaissance expert k=16</div>	Log-vraisemblance: = reco aléatoire par catégorie MRR : > reco aléatoire par catégorie	Non quantifiée dans une étude Difficulté: règles d'associations exclusions trop nombreuses

Performance moindre:

- *Positionnement des aliments (ordre et non distances absolues)*
 - *On bénéficie moins de l'historique*
- ← ++ Nombre de combinaisons d'aliments
++ Variété des aliments consommés
Séquentialité dans le repas

Modèle '*GenSeqRNN*': exploiter la séquentialité du déjeuner/diner

Entrée
Plat
Dessert

1 Axe 1:

Recommandation basée sur les
préférences utilisateurs par
apprentissage automatique

- ❑ Approche type filtrage collaboratif: '*FiltreCollab*'

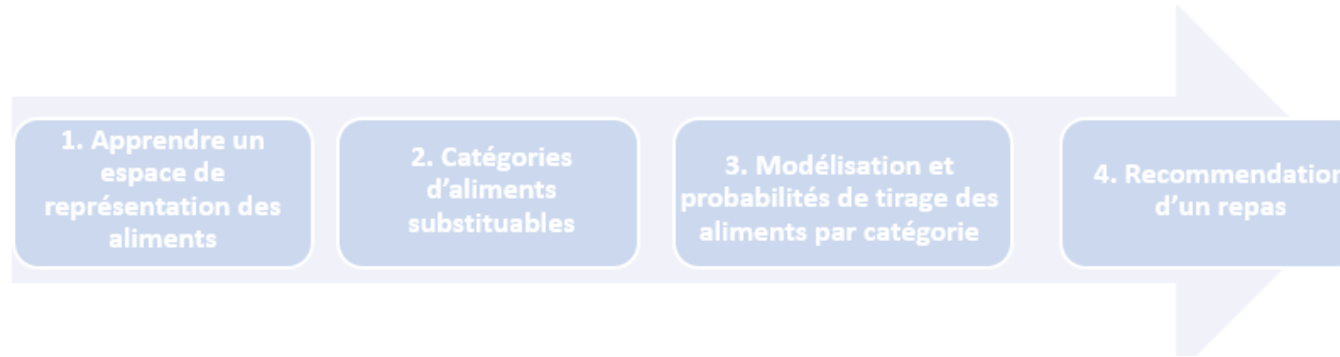


- ❑ Génération de séquences par Réseau de neurones récurrents: '*GenSeqRNN*'



Modèle '*GenSeqRNN*': testé sur le déjeuner/diner: méthodologie

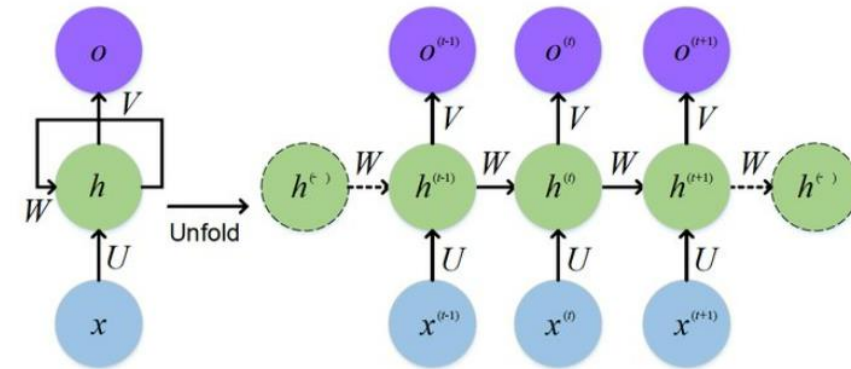
'FiltreCollab'



'GenSeqRNN'

1. Recommandation d'un repas

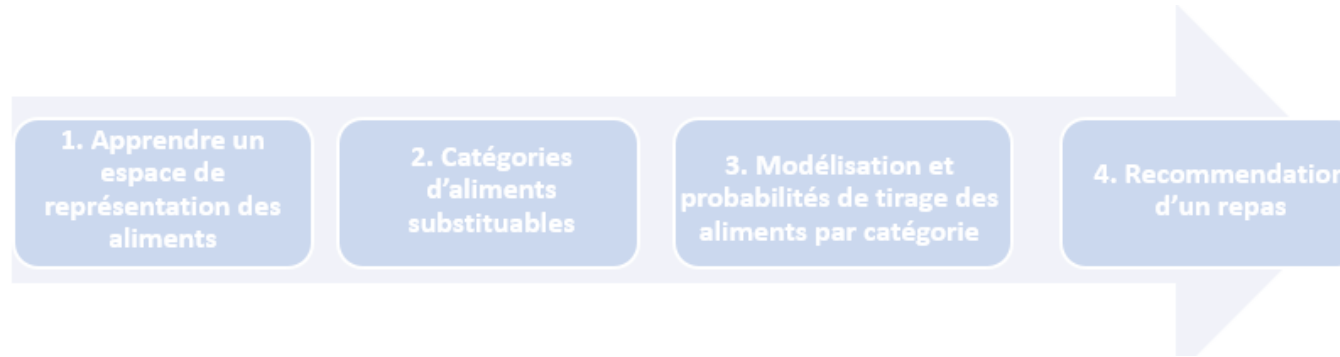
Réseau de neurones récurrents (RNN)
entraîné sur les repas
et utilisé pour générer des séquences de repas



Architecture d'un RNN

Modèle '*GenSeqRNN*': testé sur le déjeuner/diner: méthodologie

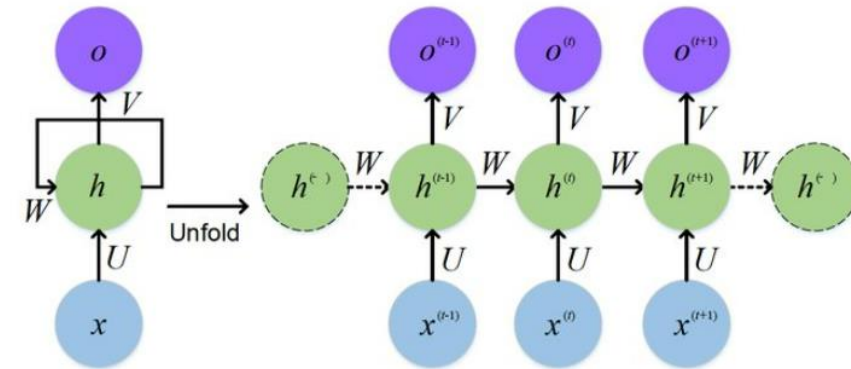
'*FiltreCollab*'



'*GenSeqRNN*'

1. Recommandation d'un repas

Réseau de neurones récurrents (RNN)
entraîné sur les repas
et utilisé pour générer des séquences de repas



Architecture d'un RNN

Critères
d'évaluation

Taux de bonne prédiction du prochain aliment: accuracy	Cohérence d'un repas (à évaluer dans 1 étude auprès d'utilisateurs)
--	--

Modèle '*GenSeqRNN*': performance sur le déjeuner/diner

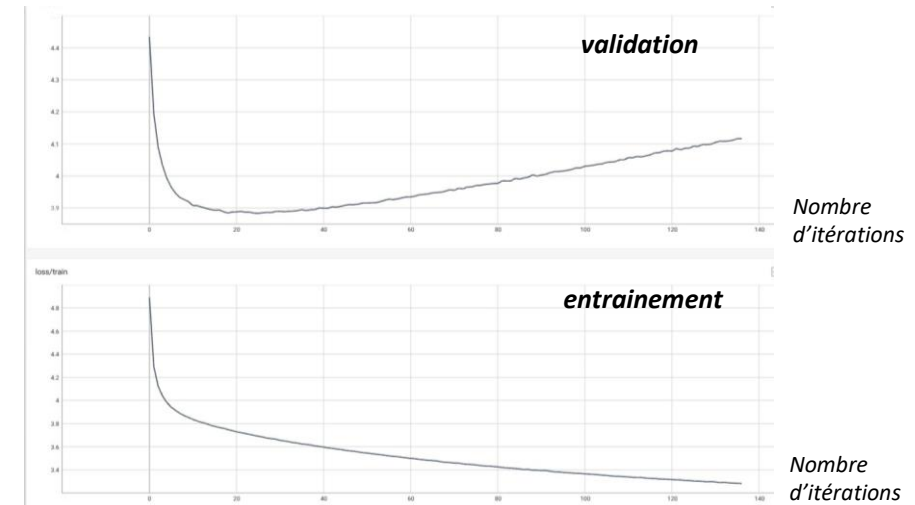
1. Recommandation d'un repas

Réseau de neurones récurrents (RNN)
entraîné sur les repas
et utilisé pour générer des séquences de repas

Hyperparamètres:

- représentations vectorielles des aliments: 10 dimensions
 - état caché: 100 dimensions, 25 itérations
 - initialisation avec les vecteurs appris par W2V
- 70% training, 15% validation, 15% test

Perte (entropie croisée)



Courbes de perte (entropie croisée) sur l'ensemble de validation et l'ensemble d'apprentissage pour le modèle RNN

Modèle '*GenSeqRNN*': performance sur le déjeuner/diner

1. Recommandation d'un repas

Réseau de neurones récurrents (RNN)
entraîné sur les repas
et utilisé pour générer des séquences de repas

Hyperparamètres:

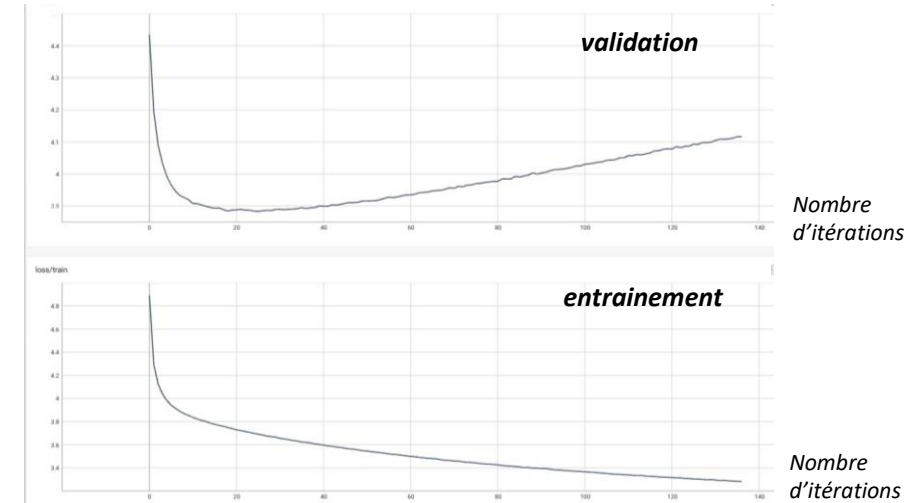
- représentations vectorielles des aliments: 10 dimensions
 - état caché: 100 dimensions, 25 itérations
 - initialisation avec les vecteurs appris par W2V
- 70% training, 15% validation, 15% test

Résultats encourageants:

Taux de bonne prédiction du prochain aliment*		Cohérence d'un repas*
Accuracy :	10%	Non quantifiée dans 1 étude utilisateurs Evaluation quali: ✓
Accuracy dans top 3 :	25%	

*Sur l'ensemble test


Perte (entropie croisée)






Courbes de perte (entropie croisée) sur l'ensemble de validation et l'ensemble d'apprentissage pour le modèle RNN

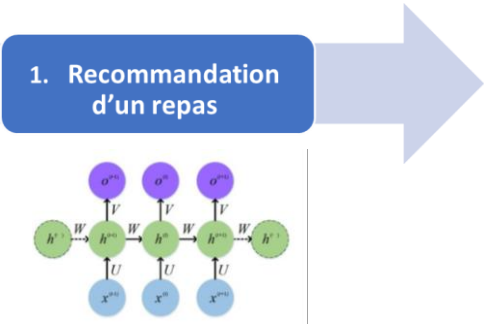
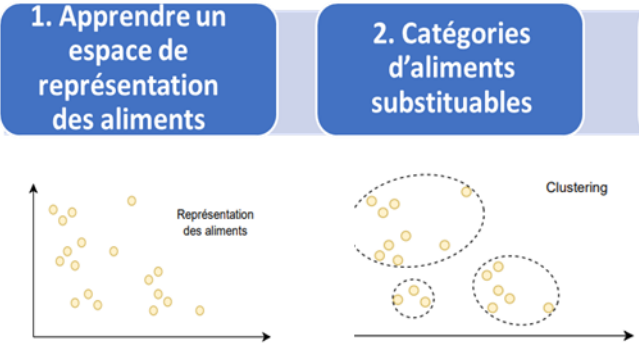


Axe 1: Comparaison des 2 modèles


	<i>Proximité des aliments apprise sur l'ensemble des utilisateurs</i>	<i>Introduction de l'utilisateur</i> 
<i>'FiltreCollab'</i>		
<i>'GenSeqRNN'</i>		

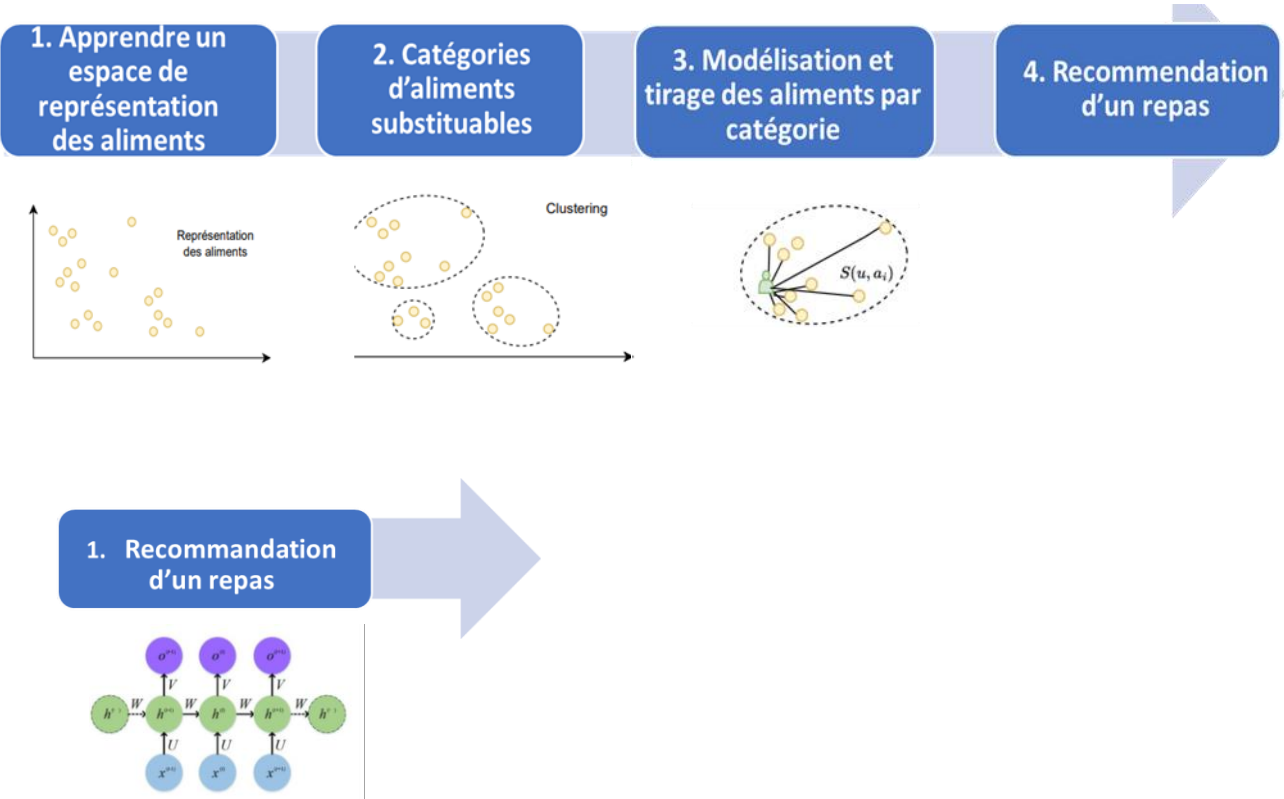
Axe 1: Comparaison des 2 modèles

	Proximité des aliments apprise sur l'ensemble des utilisateurs	Introduction de l'utilisateur 
'FiltreCollab'		
'GenSeqRNN'		



Axe 1: Comparaison des 2 modèles

	Proximité des aliments apprise sur l'ensemble des utilisateurs	Introduction de l'utilisateur 
'FiltreCollab'	✓	✓
'GenSeqRNN'	✓	✗



Axe 1: forces et faiblesses des 2 modèles

1

Axe 1:



Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

❑ **'FiltreCollab'**

❑ **'GenSeqRNN'**




Forces

- Bonne performance sur repas « répétitif » 
- **Personnalisé**
- Prometteur sur **déjeuner/diner**
- Exploite la **séquentialité** 
- **Cohérence du repas**



Faiblesses

- **Echec** sur repas plus **varié** 
- **Non personnalisé**

Axe 1: forces et faiblesses des 2 modèles

1

Axe 1:



Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

❑ **'FiltreCollab'**

❑ **'GenSeqRNN'**




Forces

- Bonne performance sur repas « répétitif » 
- **Personnalisé**
- Prometteur sur **déjeuner/diner**
- Exploite la **séquentialité** 
- **Cohérence** du repas



Faiblesses

- **Echec** sur repas plus **varié** 
- **Non personnalisé**



Pistes

- Plus de données?? (INCA2 + INCA3)
- Personnalisation de la recommandation

Connexion avec l'axe 2

1

Axe 1:

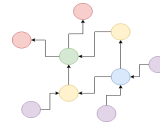
Recommandation basée sur les **préférences utilisateurs** par **apprentissage automatique**

- ❑ Approche type filtrage collaboratif: '**FiltreCollab**'
- ❑ Génération de séquences par Réseau de neurones récurrents: '**GenSeqRNN**'

2

Axe 2:

Intégration de **contraintes nutritionnelles** grâce aux **graphes de connaissance**



Génération de recommandation

✓ Tests de faisabilité
(allergies, végétarisme)



Force

- **Intégration des contraintes nutritionnelles**

Conclusion

Résultats:

Différents modèles

- Basés sur une **approche commune**
- Répondant à **différents enjeux** des systèmes de recommandation/en nutrition

➡ Défi:

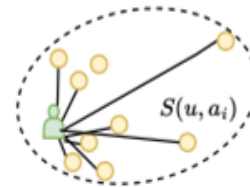
Combiner/adapter les méthodes pour **1 solution globale** de recommandation de repas

Proximité des aliments



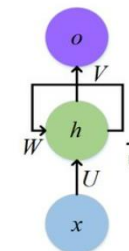
Personnalisation

'FiltreCollab'



Séquentialité

'GenSeqRNN'



Contraintes nutritionnelles spécifiques

Graphes

