

Le Traitement Automatique des Langues au service du vin

#GoûtDuVin

#IntelligenceArtificielle

#NLP

#BigData

25 Février 2021

Antoine Gérard • Winespace

Benoît Sagot • ALMAAnaCH, Inria Paris

Sylvain Thibaud • Winespace

Julien Laithier • Winespace

Déroulé

I.

INTRODUCTION

Winespace
ALMAAnaCH

II.

COLLABORATION INRIA

Projet
Méthodes
Résultats

III.

CONCLUSION

Perspectives
Bilan



I. Introduction

Winespace



Structure

Start-up Bordelaise

Activité

Solutions d'aide à la vente pour les vendeurs de vins

Expertise

Expertise "vin"

→ œnologie, sommellerie, marketing et commerce du vin, ...

Expertise "scientifique"

→ informatique, intelligence artificielle, ...

Actualité

Solution implantée dans plusieurs magasins de la région Bordelaise (Intermarché)

Importants travaux de R&D

INRIA • ALMA_{na}CH



Structure

Équipe de recherche INRIA

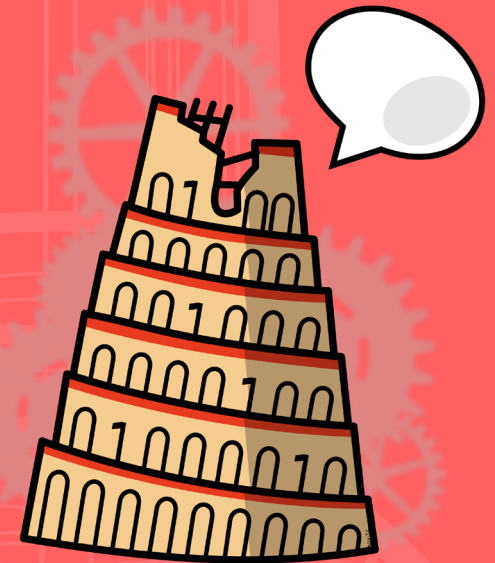
Activité

Recherche dans le Traitement
Automatique de la Langue (TAL)

Expertise

Analyse syntaxique et sémantique du
langage naturel

- Méthodes symboliques
- Méthodes statistiques
- Méthodes neuronales
- Méthodes hybrides



Collaborer avec INRIA

Expertise / Conseil

Quelques jours



Contrat de transfert

3 à 6 mois



Contrat de recherche

6 mois à 2 ans



Projet collaboratif bilatéral

2 à 3 ans

Projet collaboratif multipartenaires

2 à 4 ans



II. Collaboration Inria

Le Projet

Problématique initiale



Robe : rouge clair, légèrement tuilé.

Nez : joli fruit type cerises, fraises des bois, très frais, très engageant.

Bouche : un vin léger, gouleyant dont la fraîcheur aromatique séduit, la finale évoque la framboise, la groseille et persiste assez longuement.

Un modèle de pinot noir d'Alsace

Commentaire de dégustation



Liste de vins similaires

Le Projet

Difficultés techniques

Hétérogénéité des commentaires de dégustation

Style d'écriture, contenu, ...

Vocabulaire très spécifique

Exemple : grenache, médoc, champignonné, terroité, ...

Absence de données annotées

Aucun commentaire de dégustation n'a été manuellement annoté pour y repérer les termes pertinents ou l'associer à des informations structurées.

Réaliser de telles annotations serait coûteux, et le nombre élevé de termes rares rendrait la tâche ardue et le volume d'annotation à produire très élevé.

Les Méthodes

Approche “basique”

Objectif

Représenter les commentaires de dégustation sous forme vectorielle

Représentation Bag Of Words



Robe : rouge clair, légèrement
tuilé.

Nez : joli fruit type cerises, fraises
des bois, très frais, très engageant.

Bouche : un vin léger, gouleyant
dont la fraîcheur aromatique
séduit, la finale évoque la
framboise, la groseille et persiste
assez longuement.

Un modèle de pinot noir d'Alsace



MOT	VALEUR
Fruit	1
Pomme	0
Syrah	0
Pinot	1
Framboise	1
...	...

Les Méthodes

Approche “basée sur une ontologie”



Robe : rouge clair, légèrement
tuilé.

Nez : joli fruit type cerises,
fraises des bois, très frais, très
engageant.

Bouche : un **vin léger**,
gouleyant dont la **fraîcheur**
aromatique séduit, la finale
évoque la framboise, la groseille
et persiste assez longuement.

Un modèle de pinot noir d'Alsace



Ontologie métier

INDICE CONCEPT	NIVEAU 1	NIVEAU 2	NIVEAU 3
1	Arôme	Type	Fruit rouge
2	Arôme	Type	Gourmand
3	Acidité	Intensité	Élevée
4	Puissance	Intensité	Faible
...

INDICE CONCEPT	VALEUR	
1	1	(Fraises des bois)
2	0	
3	1	(Vin léger)
4	1	(Fraîcheur aromatique)
...

Les Méthodes

Représentation vectorielle d'un commentaire

PIPELINE

LEMMATISATION

**du commentaire de
dégustation avec MElt**

Robe : rouge clair , légèrement tuer
. Nez : joli fruit type cerise , fraise
des bois , très frais , très engageant .
Bouche : un vin léger , gouleyant dont
le fraîcheur aromatique séduire , le final
évoquer le framboise , le groseille et
persister assez longuement . Un modèle
de pinot noir de Alsace



EXTRACTIONS DES CONCEPTS

**à l'aide d'expressions
régulières**



CONSTRUCTION D'UN VECTEUR PROFIL

en utilisant l'ontologie

Encodage 1 parmi n (one-hot)
des concepts présents dans le
commentaire.



Le Projet

Recommandation de vins similaires

PIPELINE

1. CONSTRUCTION

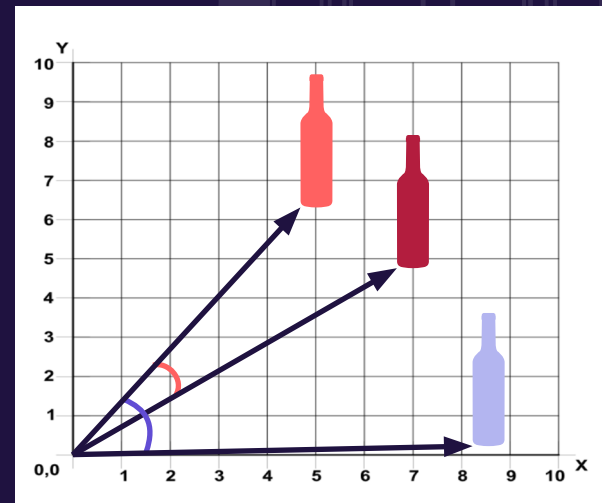
d'une matrice de "profils"

Vin	C_1	C_2	C_3	C_4	C_5	C_6	C7	C_8	C_9
1	1	0	0	1	0	0	1	0	1
2	0	1	1	0	1	1	0	1	0
3	1	1	0	1	0	0	0	0	1



2. CALCUL

de la mesure de similarité Cosinus



Les Résultats

Exemples de recommandation

Vin choisi

	ID	Couleur	Appellation	Domaine	Region	Millésime	Sucrosité
7401	GG-11082	rouge	Bordeaux supérieur	Château Bourdicotte et Grand Ferrand	bordeaux	1997	sec

Millésime: 1997, Dégustation: 1998, Différence: 1

Commentaire du vin choisi:

Robe : soutenue, jeune. Nez : fruits rouges et noirs mûrs nets. Bouche : charnue, polie, ronde, bonne tenue d'ensemble. Arômes un peu réservés. un vin franc, assez austère

Concepts du vin choisi:

arome intensite faible: arôme réserver

arome type fruit noir: fruit noir

arome type fruit rouge: fruit rouge

couleur: vin rouge

effervescence: tranquille

robe evolution faible: robe jeune

robe intensite fort: robe soutenir

rondeur intensite fort: bouche rond

sucrosité: nulle

Similarité cosinus sans acp: 0.899

Poids région: 1.000

	ID	Couleur	Appellation	Domaine	Region	Millésime	Sucrosité
121878	GG-211222	rouge	Bordeaux	Château Les Bertrands	bordeaux	2014	sec

Millésime: 2014, Dégustation: 2020, Différence: 6

Commentaire du vin similaire:

Robe concentrée, jeune. Nez timide, discrètement petits fruits rouges et noirs. Bouche lisse, bâtie sur une ligne fluide où la fraîcheur prospère. Le fruit reste discret et simple. Un bordeaux d'approche facile.

arome intensite faible: nez timide

arome type fruit noir: fruit noir

arome type fruit rouge: fruit rouge

couleur: vin rouge

effervescence: tranquille

robe evolution faible: robe jeune

robe intensite fort: robe concentrer

sucrosité: nulle

Les Résultats

Exemples de recommandation

Vin choisi

	ID	Couleur	Appellation	Domaine	Region	Millésime	Sucrosité
7401	GG-11082	rouge	Bordeaux supérieur	Château Bourdicotte et Grand Ferrand	bordeaux	1997	sec

Millésime: 1997, Dégustation: 1998, Différence: 1

Commentaire du vin choisi:

Robe : soutenue, jeune. Nez : fruits rouges et noirs mûrs nets. Bouche : charnue, polie, ronde, bonne tenue d'ensemble. Arômes un peu réservés. un vin franc, assez austère

Concepts du vin choisi:

arome intensite faible; arôme réserver

arome type fruit noir; fruit noir

arome type fruit rouge; fruit rouge

couleur: vin rouge

effervescence: tranquille

robe evolution faible; robe jeune

robe intensite fort; robe soutenir

rondeur intensite fort; bouche rond

sucrosité: nulle

Similarité cosinus sans acp: 0.874

Poids région: 0.977

	ID	Couleur	Appellation	Domaine	Region	Millésime	Sucrosité
52259	GG-105836	rouge	Côtes du Rhône	Domaine de Deurre	rhône	2003	sec

Millésime: 2003, Dégustation: 2004, Différence: 1

Commentaire du vin similaire:

Robe d'aspect jeune, pourpre violacé. Nez discret, s'ouvrant sur des notes de fruits rouges et noirs. Bouche jeune, dotée d'une matière élégante qui porte le fruit. Les tanins à caractère épicé sont plus présents en finale. Un vin équilibré.

arome intensite faible; nez discret

arome type fruit noir; fruit noir

arome type fruit rouge; fruit rouge

couleur: vin rouge

effervescence: tranquille

robe evolution faible; robe jeune

sucrosité: nulle

Les Résultats

Moyennes par région

Base de données de vin

Que voulez vous faire ?

Moyenne des concepts

Choix pour la colonne Couleur 1:

rouge

Choix pour la colonne Region 1:

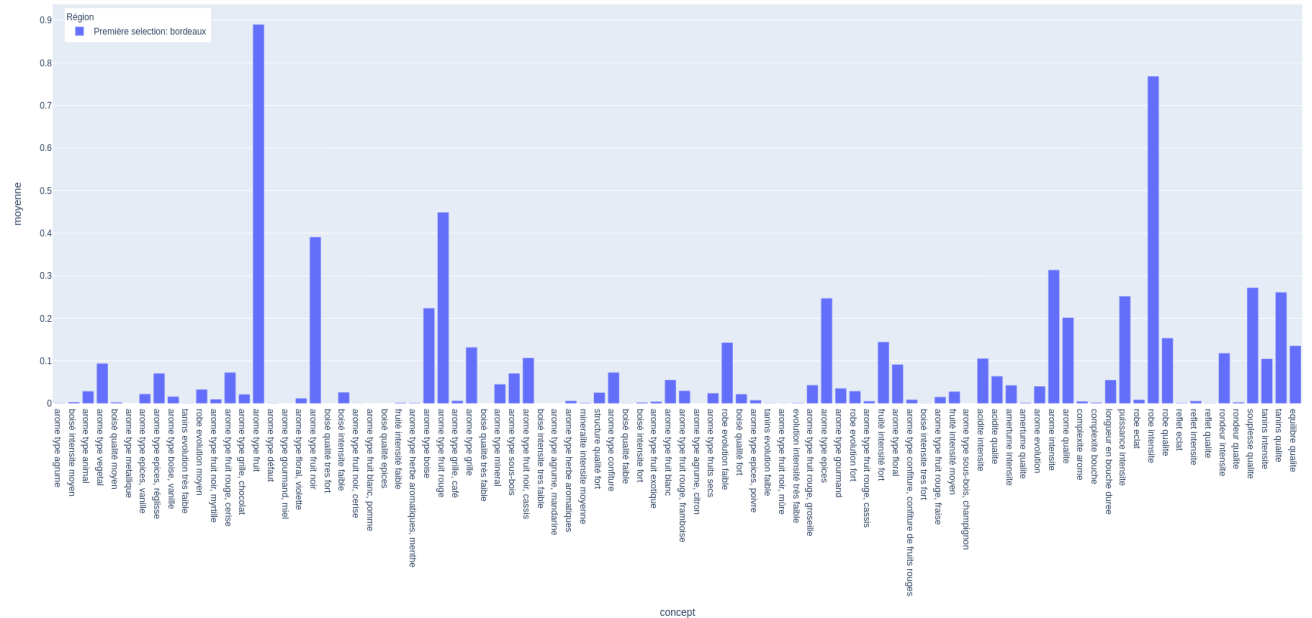
bordeaux

☒ Tout sélectionner Appellation 1

☒ Tout sélectionner Millésime 1

☐ Comparer

☐ Filtrer cépage



Moyennes par région



Que voulez vous faire ?

Moyenne des concepts

Choix pour la colonne Couleur 1:

Choix pour la colonne Region 1:

alsace ✕

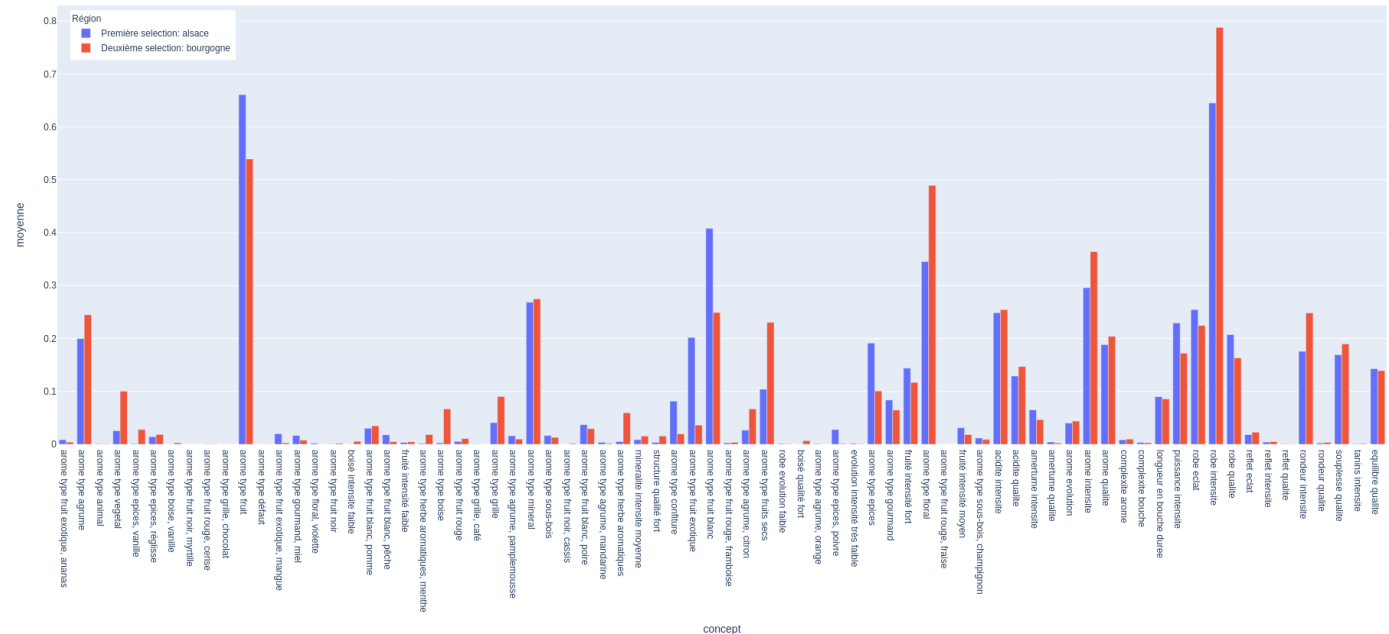
- ☒ Tout selectionner Appellation 1
- ☒ Tout selectionner Millésime 1
- ☒ Comparer

Choix pour la colonne Couleur 2:

blanc x

Choix pour la colonne Region 2:

bourgogne **x**



Les Avantages • Les Inconvénients

Avantages

- ✓ Méthode symbolique permettant une interprétation simple des résultats
- ✓ Méthode rapide à mettre en place

Inconvénients

Repose sur la recherche d'expressions régulières :

- ✗ Difficile à automatiser
- ✗ Beaucoup de regex à écrire
- ✗ Captation des nuances compliquée (adverbes, adjectifs qualificatifs, ...)

Repose sur des fichiers CSV



III. Perspectives & Conclusion

Les Perspectives

25 février 21

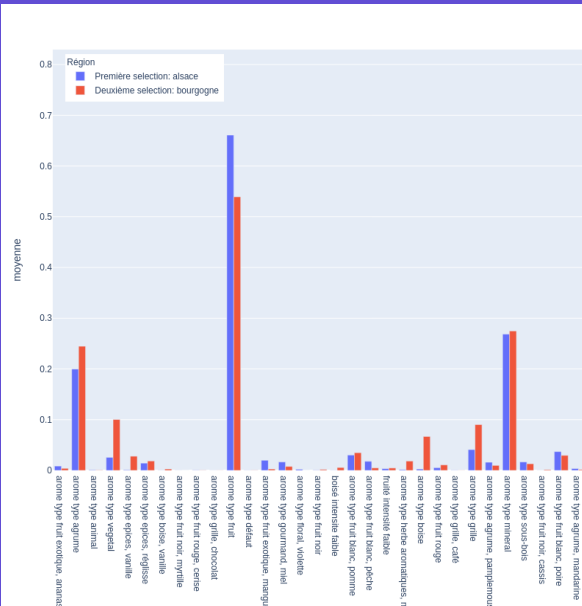
Preuve de concept / Faisabilité



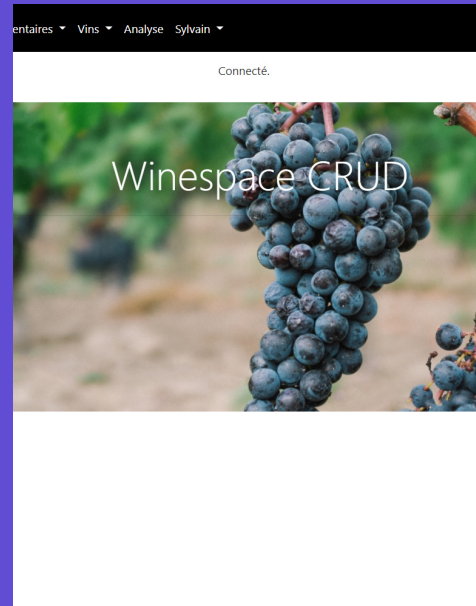
Version Bêta



Collaboration INRIA



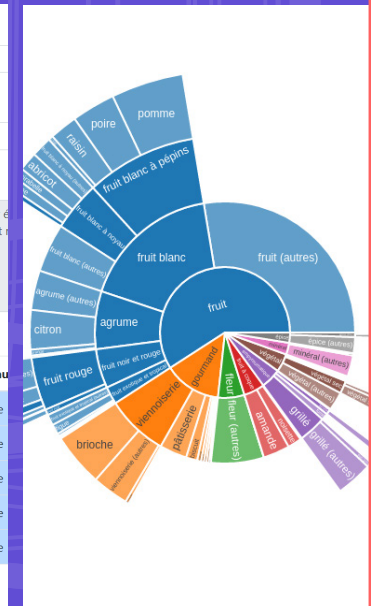
Création BDD + Interface de CRUD



Méthode d'extraction V2



Analyse + Exploitation des résultats



Les Perspectives

Analyser, comparer, comprendre
le goût des vins

**IDENTIFIER LE PROFIL TYPE D'UN
GROUPE DE VINS**

**COMPARER DES VINS OU DES
GROUPE DE VINS**

**IDENTIFIER DES SIMILARITÉS /
RAPPROCHER DES VINS**

Prédire et anticiper
le goût des vins

**ANTICIPER L'ÉVOLUTION DU GOÛT
D'UN VIN AU FIL DU TEMPS**

**PRÉDIRE LE GOÛT D'UN VIN SELON
SES CARACTÉRISTIQUES TECHNIQUES**

APPLICATIONS BUSINESS

**Recommandation
de vin / Conseil
consommateur**

Marketing Push

Sourcing Produit

**Analyse/Étude
de marché**

Bilan

Une collaboration fructueuse ...

- ✓ Qui a permis de montrer que le TAL pouvait servir le monde du vin
- ✓ Un vrai tremplin pour la R&D de Winespace
- ✓ Une preuve de concept prometteuse

... et un projet qui continue chez Winespace

- ✓ Mise en place d'une architecture solide
- ⌚ Nouvelle méthode d'extraction en cours
- ⌚ Méthode d'analyse
- ⌚ **Ouverture de la bêta-test** (sur la partie d'analyse des arômes)

Merci.



Vous souhaitez devenir bêta-testeur ?

**Vous disposez de commentaires
de dégustation ?**

Vous avez des questions ?

→ antoine.gerard@winespace.fr



Tester le conseiller digital Wia

→ <http://monconseillervin.com>

(conçu pour smartphone)

Service transfert INRIA

stip-bso@inria.fr

Site ALMA^{na}CH

<https://files.inria.fr/almanach>

Bibliothèques et cadriciels cités

