



Classifier chains pour le codage LOINC

T. MICHEL-PICQUE, S. BRINGAY, P. PONCELET, N. PATEL, G. MAYORAL

ADVANSE, LIRMM UMR 5056, Onaos





Introduction

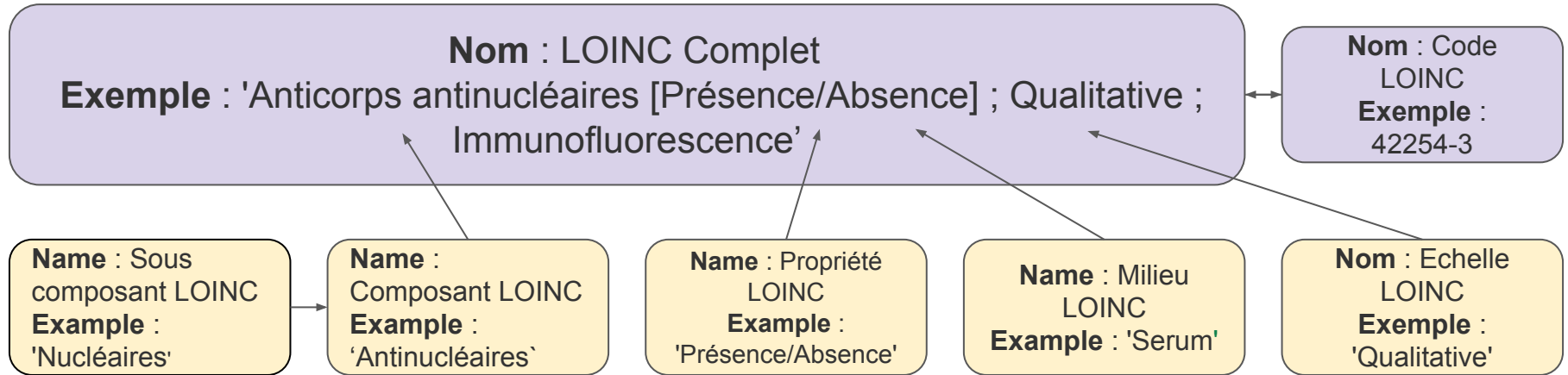
Le LOINC:

- Identifier et étiqueter les observations cliniques et les résultats de tests médicaux.

En France:

- **Depuis 2016** : Officiellement utilisée pour la biologie médicale.
- Depuis 2021 : Ségur du Numérique piloté par l'Agence du Numérique en Santé (ANS).

Code LOINC





Travaux existants

Outils de Codage Manuels/Semi-Manuels

Approche à base de règles

Avancées en Apprentissage Automatique

- Apprentissage classique
- Classification de phrases

Non prise en compte des corrélations entre composantes : méthode de chaînes classifiées inspirée de [Read et al. \(2011\)](#).



Description des données

- 139,235 entrées provenant de 99 laboratoires français, couvrant 2,463 codes LOINC

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIR E	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l



Baseline

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIR E	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l



Baseline

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIR E	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Nom : Code
LOINC
Exemple :
42254-3



Text-Based

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIR E	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Text-Based

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIR E	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Nom : LOINC Complet

Exemple : 'Anticorps antinucléaires [Présence/Absence] ; Qualitatif ;
Immunofluorescence'



Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l



Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC

Example :
'Présence/Absence'



Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC
Example :
'Présence/Absence'

Name : Milieu
LOINC
Example : 'Serum'

Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC
Exemple :
'Présence/Absence'

Name : Milieu
LOINC
Exemple : 'Serum'

Nom : Echelle
LOINC
Exemple :
'Qualitative'

Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC
Example :
'Présence/Absence'

Name : Milieu
LOINC
Example : 'Serum'

Nom : Echelle
LOINC
Exemple :
'Qualitative'

Name : Sous
composant LOINC
Example :
'Nucléaires'

Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC
Example :
'Présence/Absence'

Name : Milieu
LOINC
Example : 'Serum'

Nom : Echelle
LOINC
Exemple :
'Qualitative'

Name : Sous
composant LOINC
Example :
'Nucléaires'

Name :
Composant LOINC
Example :
'Antinucléaires'

Classifier Chains

Code client	Libellé local	Synonymes	NABM	Echelle
AC	AC ANTINUCLÉAIRE	Auto-anticorps antinucléaires	324	Numérique
Technique	Code chapitre	Système	Tube échantillon	Unités
Frottis	Calcul	Sérum	M4RT	g/l

Name : Propriété
LOINC
Exemple :
'Présence/Absence'

Name : Milieu
LOINC
Exemple : 'Serum'

Nom : Echelle
LOINC
Exemple :
'Qualitative'

Name : Sous
composant LOINC
Exemple :
'Nucléaires'

Name :
Composant LOINC
Exemple :
'Antinucléaires'

Nom : Code
LOINC
Exemple :
42254-3



Expérimentations

Baseline:

- *OneHotEncoder* + *LabelEncoder*.
- *SGD* et *XGBoost*.

Text-Based:

- *CamemBERT* et *DrBERT*.
- 20 epochs, batch de 64, optimiseur *AdamW*.

Classifier Chains:

- *OneHotEncoder* + *LabelEncoder*.
- *SGD* et *XGBoost*.



Résultats

Méthode	Précision (Pondérée) (%)	Rappel (Pondéré) (%)	Score F1 (Pondéré) (%)
Baseline SGD	0,85	0,83	0,84
Baseline XGboost	0,85	0,75	0,80
Text-based CamemBERT	0,84	0,80	0,82
Text-based DrBERT	0,81	0,81	0,81
Classifier chains SGD	0,88 (+0,03)	0,87 (+0,04)	0,87 (+0,03)
Classifier chains XGBoost	0,81 (-0,04)	0,82 (+0,07)	0,82 (+0,02)



Conclusions

Comparaison des Approches:

- Expériences approfondies sur le codage LOINC en français.
- Approches testées : apprentissage automatique, modèles linguistiques, classifieur chaînes.
- Classifieur chaînes avec **SGD** atteint un score F1 de 0,87.

Prochaines Étapes:

- Tester sur divers ensembles de données, incluant le jeu de données MIMIC.
- Expériences en combinant l'approche classifieur chaînes et modèles linguistiques.