

Classement du fromage cheddar Par une approche prédictive

A. Temimi, A. Engel et J. Chamberland

STELA-Centre de recherche en sciences et technologie du lait, INAF, Département des sciences des aliments, Université Laval
Institut Agro Dijon, Dijon, France

Contexte

Le cheddar est classé selon :

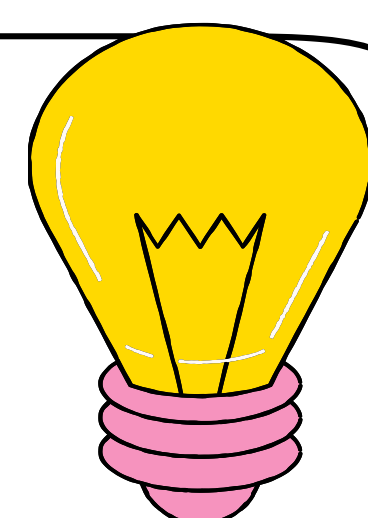
- Le degré de maturation (doux, moyen, fort)
- Le format de conversion (bloc ou tranche)

Le classement des fromages est basé sur des techniques subjectives, humaines.



Objectifs

Créer un modèle prédictif facilitant le processus de classement des fromages.



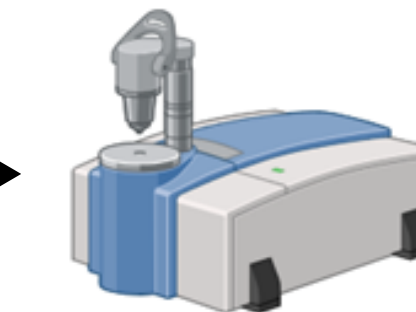
Méthodologie

Collecte de cheddars industriels (n = 144)

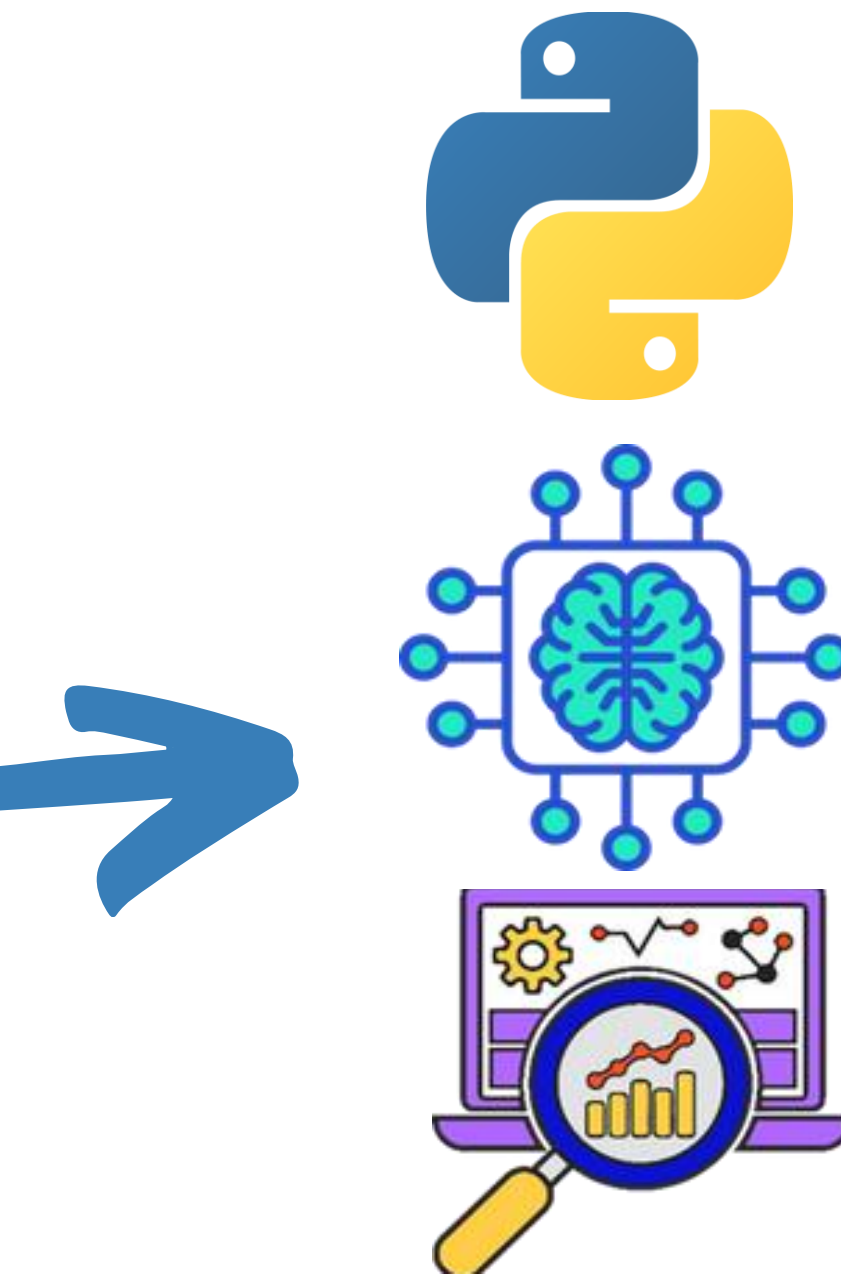


Génération des données :

- Méthodes spectroscopiques : MIR et UV-Vis-NIR
- Méthode chromatographique : Flash-GC
- Profil de texture
- Caractérisation physico-chimique



Notes de classement des fromages

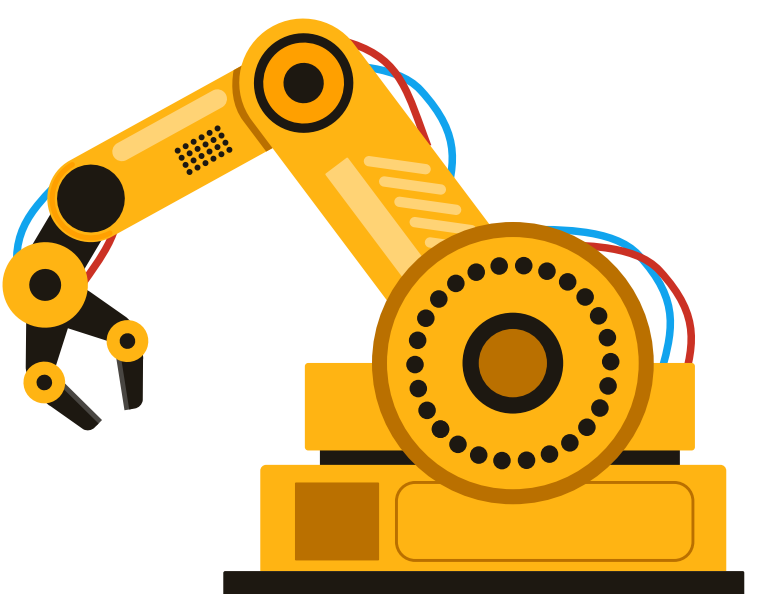


Traitement des données et création des modèles prédictifs

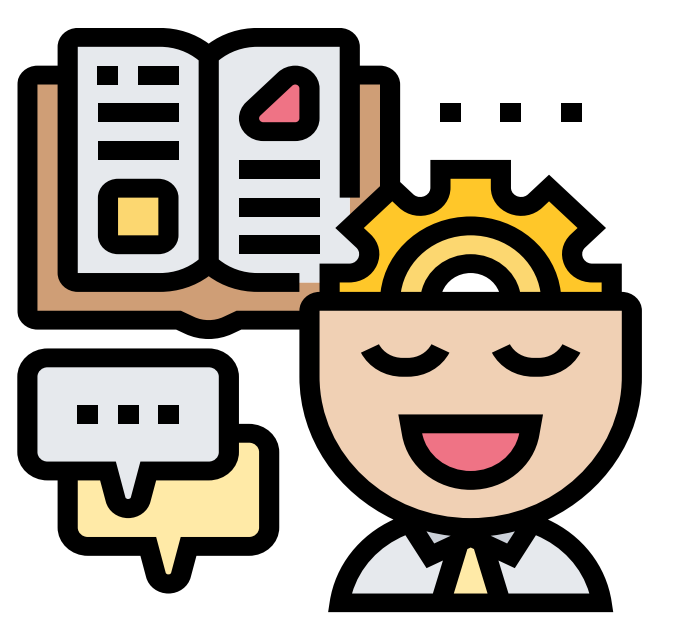
Applications et perspectives



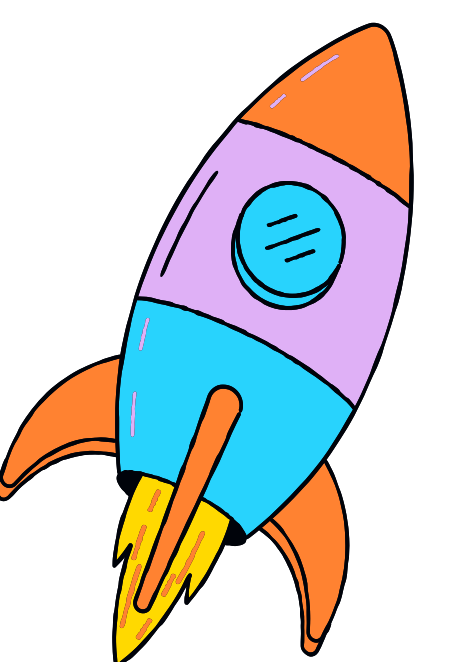
Compléter le jugement humain sans le remplacer



Automatiser les processus en usine



Améliorer les modèles de prédiction

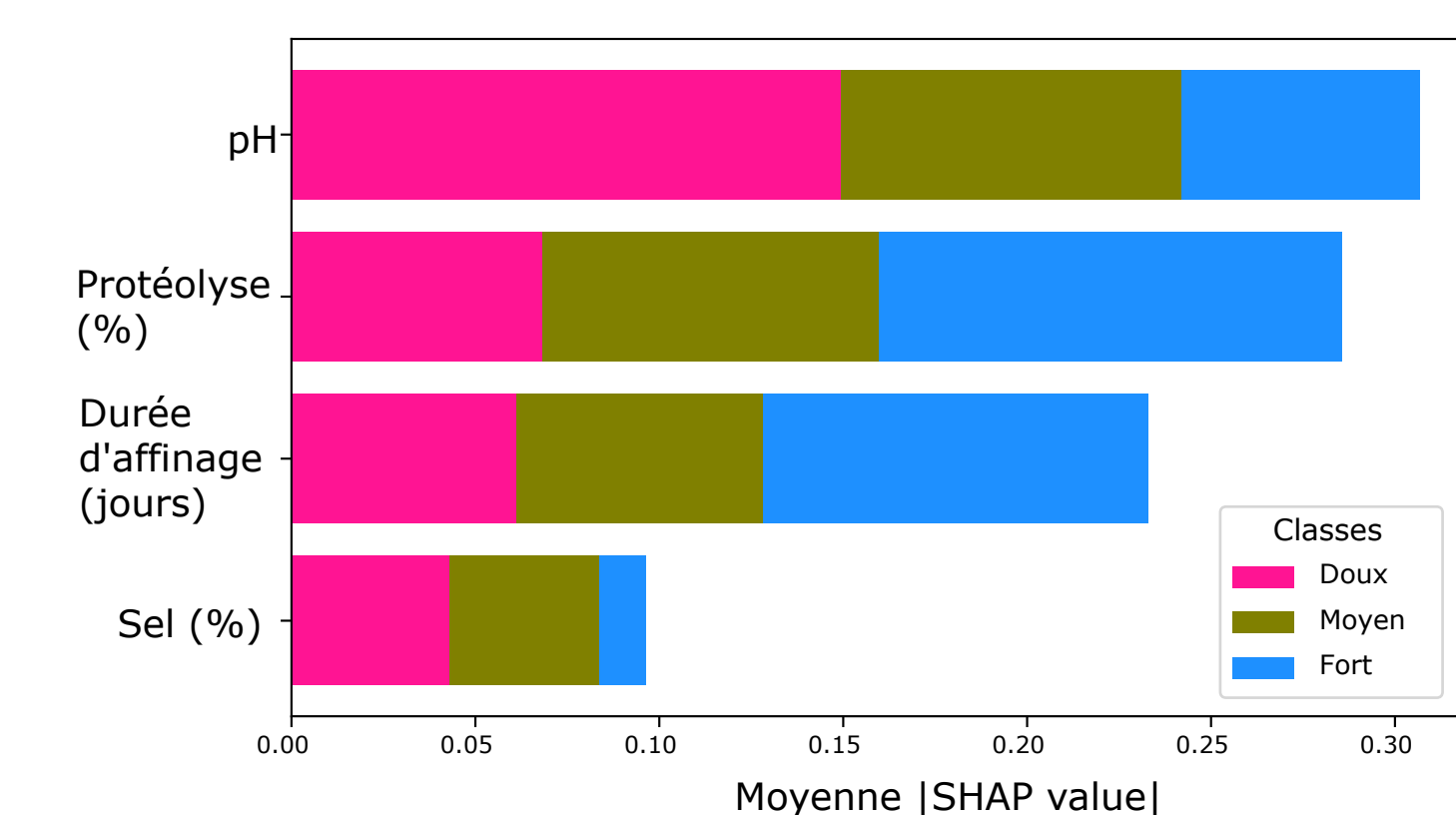


Explorer des approches avancées

Résultats : Degré de maturation

1 Jeu de données physico-chimiques

Variables explicatives



Classe réelle	Doux	Moyen	Fort
Doux	6	2	1
Moyen	1	9	0
Fort	0	0	10

Matrice de confusion

*K= Cohen's kappa

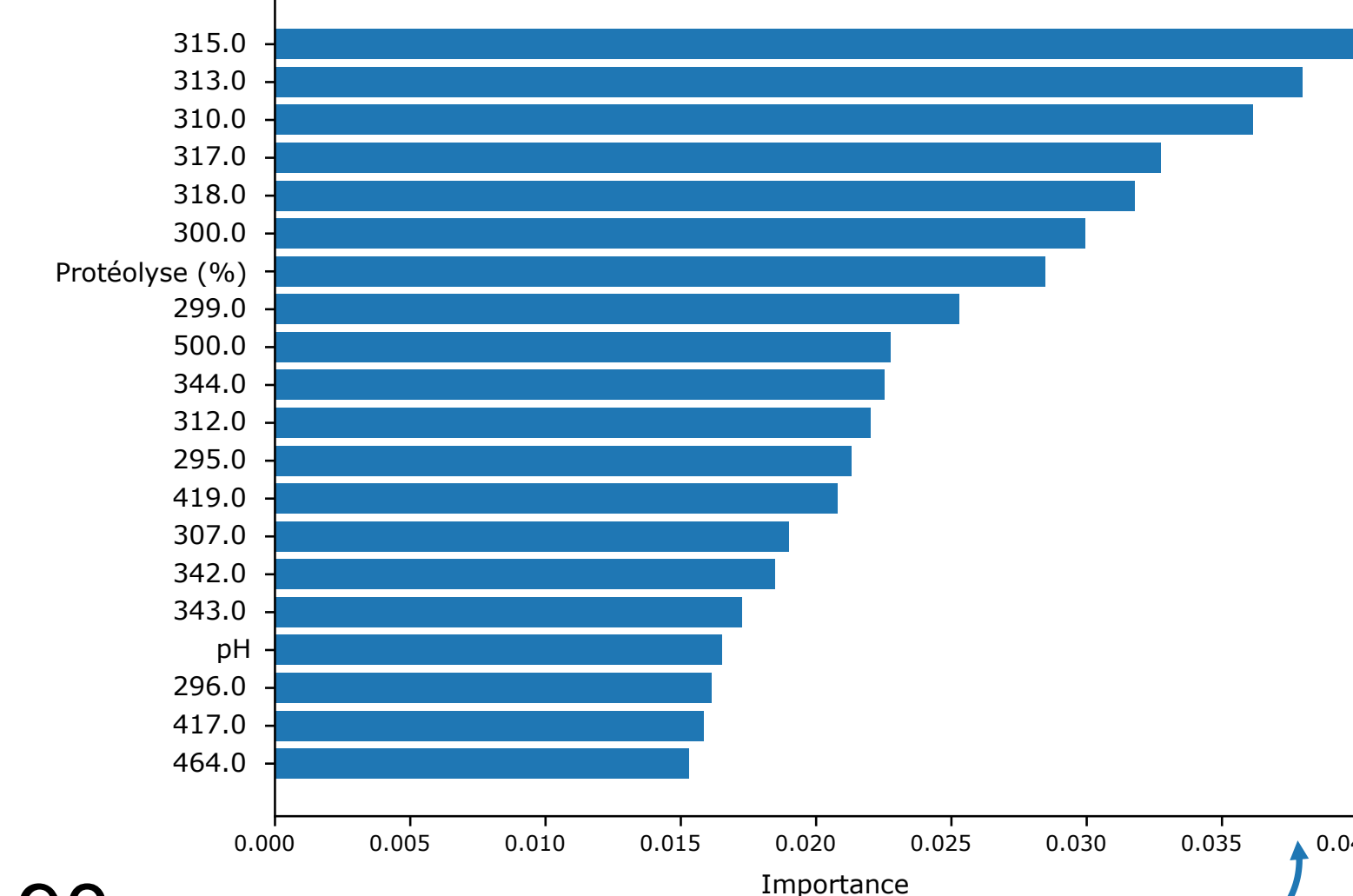
2 Jeu de données spectrales (UV-Vis-NIR) + physico-chimiques

Matrice de confusion

Classe réelle	Doux	Moyen	Fort
Doux	9	0	0
Moyen	2	8	0
Fort	0	0	10

Exactitude : 93% ; K = 0,90

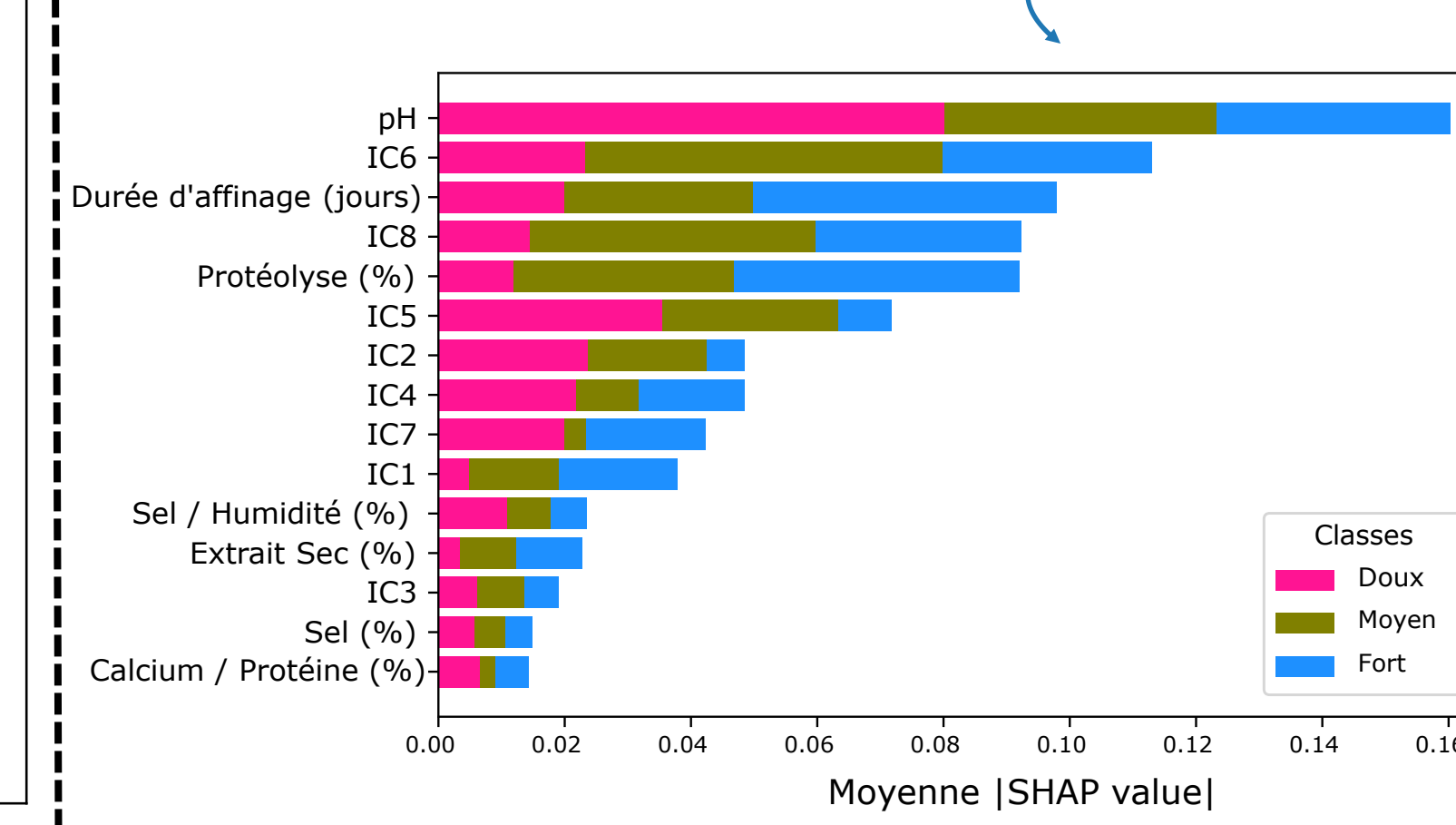
Modèle le + performant : XGBoost



Variables explicatives

3 Jeu de données spectrales réduites (UV-Vis-NIR/ICA) + physico-chimiques

Variables explicatives



Classe réelle	Doux	Moyen	Fort
Doux	9	0	0
Moyen	1	9	0
Fort	0	0	10

Matrice de confusion

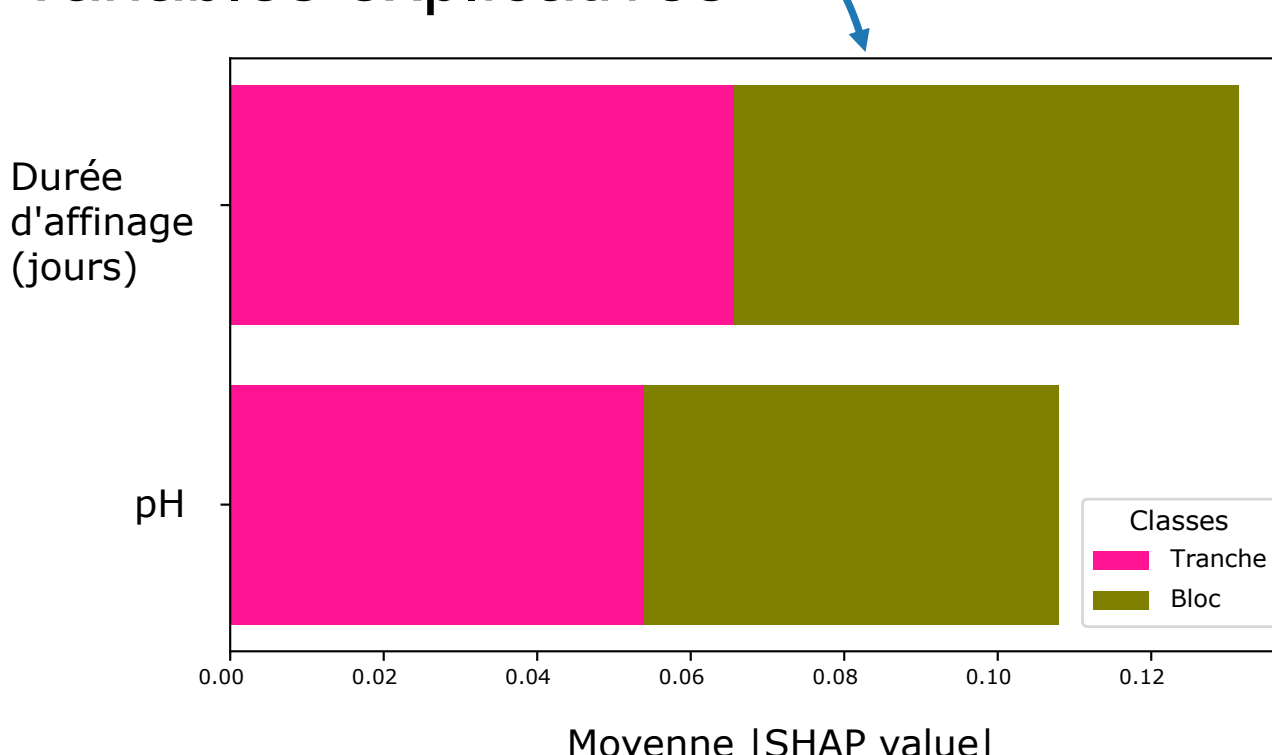
Exactitude : 97% ; K = 0,95

Modèle le + performant : AdaBoost

Résultats : Types de conversion

1 Jeu de données physico-chimiques

Variables explicatives



Classe réelle	Tranche	Bloc
Tranche	9	5
Bloc	4	11

Matrice de confusion

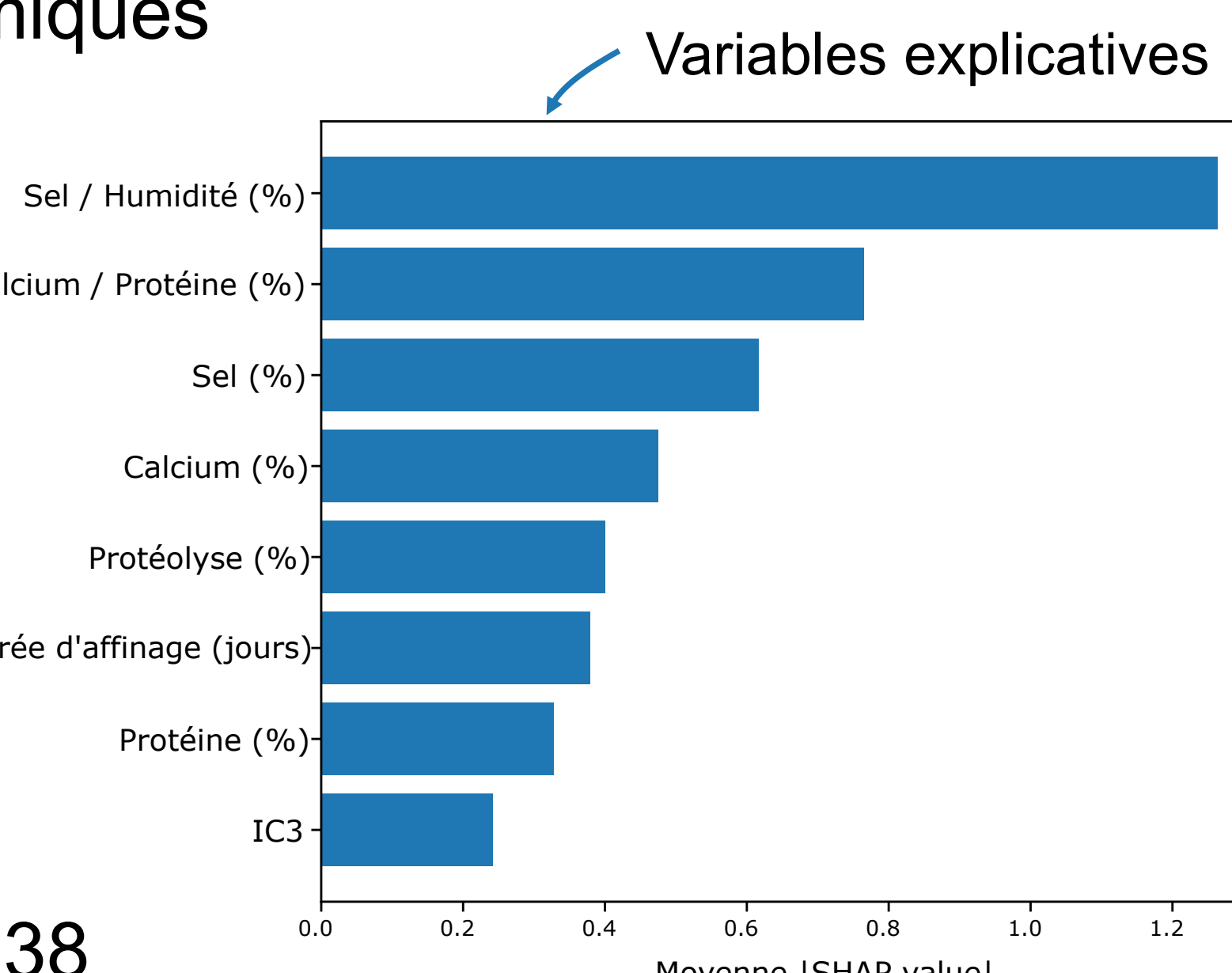
2 Jeu de données spectrales réduites (MIR/ICA) + physico-chimiques

Matrice de confusion

Classe réelle	Tranche	Bloc
Tranche	9	5
Bloc	4	11

Exactitude : 69% ; K = 0,38

Modèle le + performant : Régression logistique



Variables explicatives

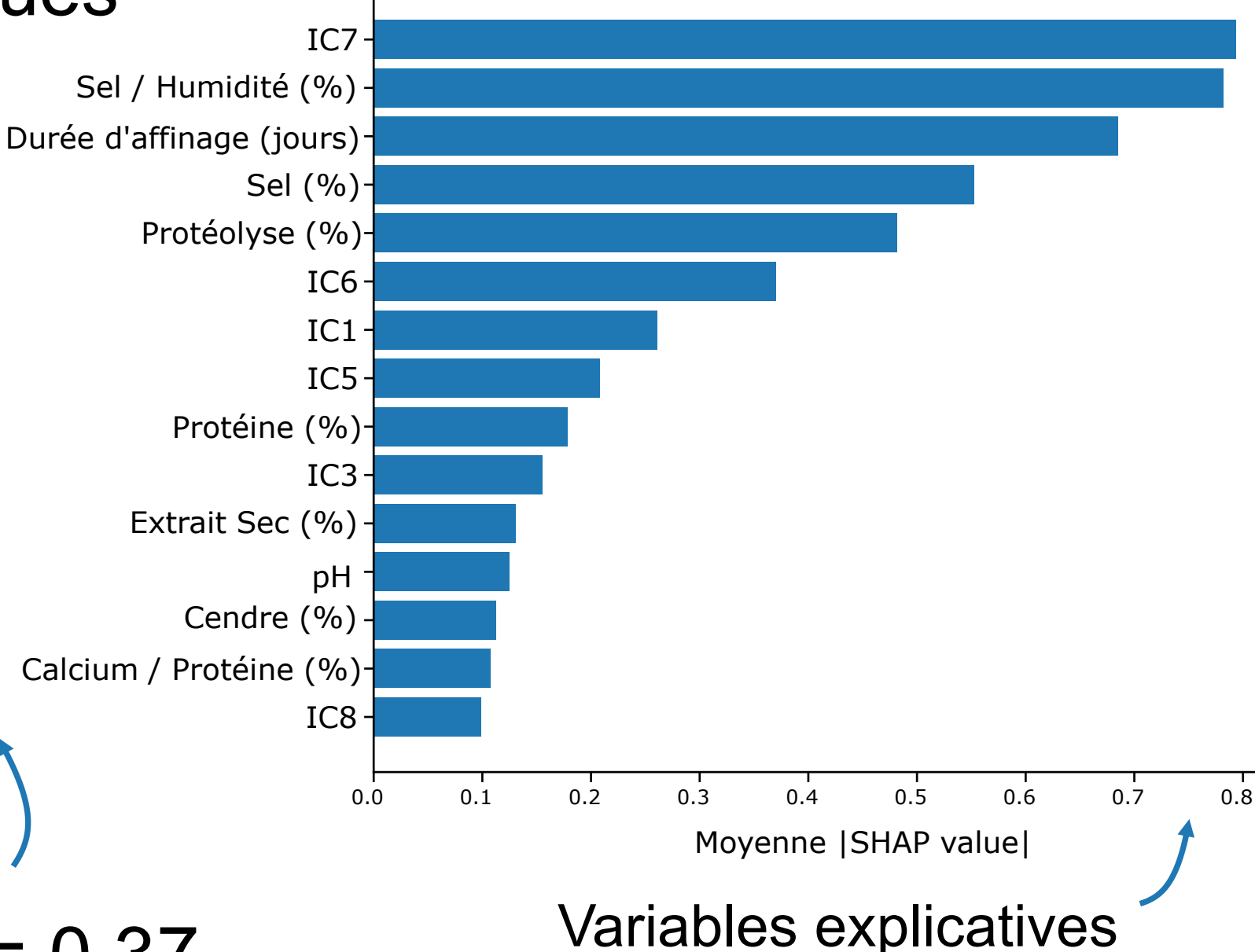
3 Jeu de données spectrales réduites (UV-Vis-NIR/ICA) + physico-chimiques

Matrice de confusion

Classe réelle	Tranche	Bloc
Tranche	8	6
Bloc	3	12

Exactitude : 69% ; K = 0,37

Modèle le + performant : Régression logistique



Variables explicatives