Conception d'un Outil de Diagnostic: Application à l'Essai de Panification en Industrie Boulangère

Melanie Munch[1], Cedric Baudrit[2], Kamal Kansou[3], Christophe Fernandez[2]

[1] I2M-Université de Bordeaux, Talence, France
[2] INRAE-I2M-Université de Bordeaux, Talence, France
[3]UMR INRAE-BIA, Nantes, France

Abstract

Comme l'ensemble du secteur agricole, la production de blé en France doit s'adapter au changement climatique et des pratiques de production, ainsi qu'à l'évolution des attentes des consommateurs. Ce contexte conduit le secteur à revoir des critères de qualité d'un blé en France et à repenser l'usage des données issues des différents tests technologiques en tirant partie des développements dans le domaine de l'IA. Un outil d'évaluation de la qualité des blés est en développement dans le cadre du projet EVAGRAIN[1] pour prédire la qualité espérée de blés, autrement dit anticiper sa valeur pour un usage donné (en panification, biscuiterie, amidonerie ou alimentation animale). Dans cet article, nous présentons le travail de modélisation en cours, décrivant une première approche de cet outil basé sur le couplage d'un modèle de raisonnement qualitatif avec un réseau Bayésien.

1 Introduction

Le dérèglement climatique, les enjeux sociétaux de durabilité, les demandes des consommateurs pour une production locale et saine sont autant d'éléments auxquels le secteur céréalier devra s'adapter. Cette adaptation passe par une révision de l'évaluation de la qualité d'un blé, véritable pierre angulaire du secteur qui détermine par exemple le prix d'un lot de blé sur le marché. Différents tests technologiques et analytiques existent pour caractériser un blé, c'est leur combinaison associée à une expertise "métier" qui permet au technologue de formuler un profil et d'envisager un usage approprié en meunerie, biscuiterie, amidonerie, etc. Le projet ANR Evagrain [2] a été défini pour répondre à cette problématique, en proposant de nouveaux critères visant à mieux comprendre le comportement du blé lors du procédé de panification (opération consistant à transformer la farine en pain) et ainsi informer sur la qualité de celui-ci. A partir de ceux-ci, un outil d'aide à la décision (DSS) permettra à terme de qualifier les blés et proposer aux utilisateurs (meuniers, boulangers) l'utilisation la plus adaptée.

Afin de construire ce DSS, il est d'abord nécessaire de modéliser le test le plus déterminant pour l'évaluation de la qualité du blé : l'essai de panification. Défini par la norme AFNOR NFV03-716, ce test décrit précisément les différentes étapes de fabrication d'un pain, du mélange des ingrédients jusqu'à la sortie du four. Tout au long de ce procédé, plusieurs mesures sensorielles (*e.g.* collant de la pâte, couleur de la mie) sont effectuées, tout d'abord sur la pâte, puis le produit fini (pain, croûte, mie). Celles-ci permettent d'évaluer différentes caractéristiques de la farine - et donc par extension du blé qui la compose. S'il est particulièrement riche et proche de l'usage finale du blé, il s'agit en revanche d'un essai coûteux qui repose sur la "sensibilité" d'un évaluateur humain.

Dans cet article, nous nous intéresserons donc d'une part aux différentes mesures sensorielles évoquées, et nous étudierons les relations pouvant exister entre celles-ci. Pour cela, nous nous appuierons sur les réseaux Bayésiens (RBs), qui ont l'avantage de pouvoir facilement représenter des réseaux complexes entre plusieurs variables teintées d'incertitude. Dans un premier temps, nous définirons nos objectifs et la structure du modèle choisi pour y arriver (Section 2). Nous présenterons les différents outils utilisés pour la construction de ce modèle (Section 3). Enfin, nous présenterons le domaine d'application et les résultats obtenus (Section 4).

L'originalité de ce travail vient à la fois (1) de la base de connaissance utilisée, résultat d'une collecte réalisée sur les 20 dernières années; (2) de l'interaction définie entre un RB appris à partir de cette base et un système expert préexistant.

2 Modélisation

L'objectif de ce travail est de proposer un modèle combinant le pouvoir de raisonnement dans l'incertain des RBs avec des connaissances expertes. L'enrichissement apporté par les connaissances expertes permettrait alors d'améliorer l'évaluation de la qualité du blé et la prédiction de son comportement lors du test de l'essai de panification.

La structure adoptée est décrite en Fig.1, et se compose en deux parties:

1. Dans un premier temps, l'utilisateur entre les caractéristiques mesurées du blé. L'effet sur les notes de pétrissage (premiers descripteurs de la pâte) est estimé à l'aide d'un modèle de raisonnement qualitatif.
2. Les notes déduites à partir des règles logiques sont reportées dans un réseau bayésien reprenant l'ensemble

[1]https://anr.fr/Projet-ANR-20-CE21-0008
[2]https://anr.fr/Projet-ANR-20-CE21-0008

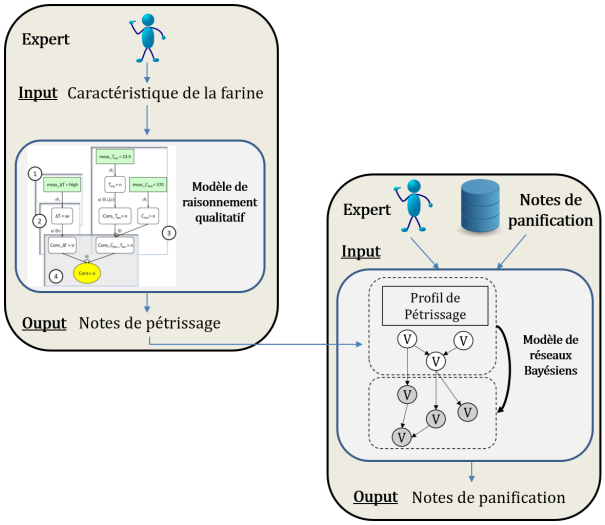

Figure 1: Structure du modèle présenté dans cet article. Les caractéristiques de la farine (quantité de protéines, amidon endommagé et pentosanes) sont connues, et permettent de définir un profil de pétrissage à partir de règles logiques issues du modèle de raisonnement qualitatif. Ce profil permet de fixer certaines variables du RB (en blanc), et d'inférer les autres (en gris).

des notes de l'essai de panification. Les autres notes sont alors inférées à partir de ces premières.

La difficulté de l'évaluation d'un essai provient de la méconnaissance des interactions entre les caractéristiques du blé, de la farine et des appréciations sensorielles réalisées pendant l'essai. Dans notre cas, le RB apprit permet d'identifier les combinaisons de notes les plus probables à partir des notes déjà observées. Dans la suite de cet article, nous nous concentrerons sur l'apprentissage du RB guidé par les connaissances expertes, ainsi que son évaluation.

3 Travaux antérieurs

Dans la section précédente nous avons pu présenter nos objectifs de modélisation. Dans cette section, nous présenterons brièvement les travaux antérieurs sur lesquels se base notre modèle développé. Le formalisme des RBs a été choisi ici pour favoriser l'explicabilité des résultats, et étudier les liens découverts entre les observations.

Modèles Probabilistes: RB et MRP

Réseaux Bayésiens Un RB est la représentation de probabilités jointes sur un ensemble de variables aléatoires dont les dépendances probabilistes sont encodées par un graphe acyclique orienté. Ainsi, chaque nœud représente une variable, et chaque flèche une relation probabiliste. L'apprentissage d'un RB se fait généralement en deux temps : la structure d'abord, puis les dépendances probabilistes. Dans notre cas, cet apprentissage se fait en utilisant l'algorithme classique Greedy Hill Climbing [1] guidé par un score BIC [11].

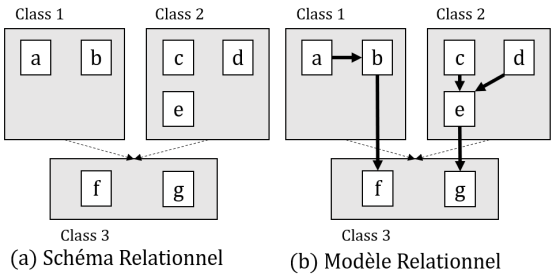

Figure 2: Exemple d'un MRP et de la définition de ses (a) hautes et (b) basses structures.

Modèles Relationels Probabilistes Le MRP est une extension orientée objet des RB. Là où un RB ne nécessite qu'une seule couche d'information pour être décrit, le MRP en nécessite deux : (1) le schéma relationnel (SR), qui fournit une description qualitative des classes et de leurs variables (appelées dans ce cadre attributs) composant sa structure ; et (2) le modèle relationnel (MR), qui contient toutes les informations quantitatives sur la distribution probabiliste entre les attributs. Deux classes peuvent être liées dans le SR par des liens relationnels, qui indiquent la direction des liens probabilistes entre les attributs : ainsi, dans la Figure 2, le lien de **Classe 1** vers **Classe 3** dans le SR force l'orientation de la relation entre les attributs b et f dans le MR. En revanche, en l'absence de lien relationnel (comme par exemple entre **Classe 1** et **Classe 2**), aucun lien ne peut être appris. Enfin, l'orientation des liens au sein d'une même classe n'est pas forcée. L'apprentissage entièrement automatique d'un MRP, en deux temps (pour le SR puis le MR) est complexe : dans notre cas, nous construisons à la main le SR, réduisant la complexité d'apprentissage du MR à celle d'un RB [6]. Une fois le SR et le MR définis, le MRP est similaire à un RB.

Apprentissage sous contraintes L'apprentissage sous contraintes (*i.e.*, guidé par des connaissances restreignant l'espace de recherche) dans le cadre des RB permet d'améliorer grandement la précision des résultats, qu'il s'agisse de celle de la structure [3] ou des paramètres [4]. Cela se vérifie d'autant mieux dans le cadre de petites bases de données [9]. Dans cette optique, des travaux précédents ont proposé des méthodes pour imposer un ordre total [2] ou partiel [10] sur les nœuds, créant ainsi des contraintes directionnelles sur leurs relations. Dans notre cas, la définition manuelle du SR permet de créer un ordre partiel entre les variables. Celui-ci transcrit les connaissances expertes apportées par la base de connaissance (BC) et l'expert, favorable à un contexte de découverte causale comme présenté dans [8].

Par définition, les relations entre classes du SR établissent les contraintes de précédence : si une telle contrainte existe entre la classe comprenant la variable A vers la classe comprenant la variable B, alors un lien appris sera forcément orienté de A vers B. Au contraire, deux variables définies dans une même classe n'auront aucune contrainte de direction quant à leur potentielle relation. Ces contraintes de

précédence traduites en liens relationnels entre classes peuvent être déduites des informations temporelles contenues dans le BC (par exemple, une mesure faite à une étape au temps t peut avoir une influence sur une mesure faite à une étape au temps $t + n$, mais pas l'inverse). Elles peuvent être également fournies par l'expert, en tant que causalités potentielles du type "Je sais que la variable A peut expliquer la variable B" ou, au contraire, "A et B ne peuvent pas partager de lien".

Dans notre cas, nous définissons un SR à partir du système décrit et des connaissances expertes (présenté en Figure 3) afin de guider l'apprentissage.

Système expert ASCOPAIN

Le système expert ASCOPAIN a été développé afin de modéliser l'étape de pétrissage lors du procédé de panification [7]. Il permet ainsi de faire le lien entre différentes caractéristiques de la farine et des "profils" de pétrissage, c'est-à-dire les notes obtenues pendant l'essai de panification lors de cette étape. Ainsi, le système est capable de nous fournir une information experte conséquente, qui permettra de guider les prédictions du RB appris.

4 Application à l'Essai de Panification

L'essai de panification est un test destiné à évaluer la qualité d'une farine donnée. Dans le cas du projet ANR Evagrain, ces farines sont composées uniquement d'un type de blé donné: par extension, le test permet donc de donner une idée de la qualité boulangère de ce blé dans le cadre du procédé de panification. Dans un premier temps, nous décrirons plus précisément les différentes étapes, ainsi que les mesures sensorielles récupérées durant ce test. Nous montrerons ensuite comment ces données ont été pré-traitées afin d'être utilisées dans le cadre d'un apprentissage de RB. Enfin, nous détaillerons les connaissances expertes récupérées auprès d'experts du domaine.

Description du Domaine

Caractérisation du Blé

Structure du Test Le test est composé de six étapes successives de mesures majoritairement sensorielles dont les comportements associés sont décrits dans la norme AFNOR:

- **Pétrissage.** Les ingrédients (principalement farine et eau) sont mélangés, afin de créer la pâte. La quantité d'eau ajoutée est représentée par la variable *Hydratation*. Six mesures sensorielles sont effectuées sur la pâte: *Consistance*, *Lissage*, *Collant de la pâte*, *Extensibilité*, *Elasticité*, *Relâchement*.
- **Pointage.** La pâte est laissée au repos. Une mesure sensorielle est réalisée: *Détente: Relâchement*.
- **Façonnage.** La pâte est divisée en plusieurs portions, appelés les pâtons. Ils sont façonnés de façon à donner la forme finale du pain. Quatre mesures sensorielles sont réalisées: *Allongement*, *Déchirement*, *Élasticité*, *Collant de la Pâte*.
- **Apprêt.** Les pâtons sont à nouveau mis au repos, et deux mesures sont effectuées: l'*Activité fermentaire* et le *Déchirement de la pâte*.
- **Mise au four.** Les pâtons sont cuits au four. Deux mesures sont réalisées juste avant l'enfournement: *Collant de la pâte* et *Tenue de la pâte*.
- **Évaluation du produit final.** Celle-ci a lieu en deux temps: tout d'abord, le pain complet est évalué à travers différentes mesures sensorielles: *Section*, *Couleur*, *Epaisseur*, *Croustillant*, *Coup de lame - Développement*, *Coup de lame - Régularité* et *Coup de lame: Déchirement*. La mie est ensuite évaluée, à travers cinq mesures sensorielles: *Couleur*, *Texture - Plasticité*, *texture - Collant*, *Alvéolage - Régularité* et *Odeur*.

Hormis le volume et l'hydratation, qui sont des mesures quantitatives, les mesures sensorielles sont évaluées sur échelle de notation d'au maximum 7 valeurs:

- **Vraiment Insuffisant** (VI), **Insuffisant** (I) et **Faiblement Insuffisant** (FI) si les propriétés observés sont insuffisantes par rapport à la normale.
- **Normal** (N) si les propriétés observées sont celles attendues et décrites par la norme.
- **Faiblement en Excès** (FE), **Excès** et **Vraiment en Excès** (VE) si les propriétés observées sont en excès par rapport à la normale.

Certaines propriétés s'observent uniquement en excès ou insuffisance. Par exemple, une pâte normale n'étant pas collante, le Collant se mesure uniquement en excès. Dans ce cas de figure l'échelle de notation est réduite. A ces mesures sensorielles qualitatives s'ajoutent également deux mesures quantitatives: *Hydratation* (quantité d'eau ajoutée au **Pétrissage**) et *Volume* (mesure du volume du pain après la cuisson).

En conclusion, nous disposons de 26 mesures sensorielles et 2 mesures quantitatives. L'un des objectifs dans cet article sera de voir s'il est possible d'apprendre un réseau capable de modéliser les liens existants entre ces différentes variables, de façon à pouvoir simuler des essais de panification.

Choix de modélisation Dans le cadre de l'ANR Evagrain, un jeu de données reprenant 10,898 essais de panifications s'étalant de 2003 à 2022 réalisés par l'institut technique Arvalis [3] a été utilisé. Il est d'abord important de noter que bien que très standardisée, l'évaluation des notes sensorielles a forcément une composante "expérimentateur" : tout le monde n'a pas la même sensibilité et, pour un même échantillon donné, différentes personnes pourraient donner différentes notes. Néanmoins, les tests présents dans le jeu de données ayant tous la même origine (l'institut technique, dont les expérimentateurs ont été formés pour produire des résultats comparables), nous formulons un premier choix de modélisation: **(C1) la variable cachée "expérimenteur" n'a pas d'influence sur les résultats présents dans le jeu de données**. Ainsi, nous considérons que même si certaines notations ne suivent pas la même logique interne, le nombre de répétitions est suffisamment important pour permettre de

[3] https://www.arvalis.fr/

considérer que notre modèle représentera les relations entre les notes de façon cohérente, au moins dans le cadre de l'essai réalisé par Arvalis.

Afin de discrétiser, nous avons traité différemment les variables numériques et les notes de l'essai de panification. Pour les premières, nous avons choisi de travailler avec une répartition en trois quantiles, afin de refléter les valeurs plutôt basses/moyennes/élevées par rapport au jeu de données. Pour les variables qualitatives, l'objectif principal n'est pas de prédire la note exacte, mais la présence ou absence de problème. Ainsi, nous posons un deuxième choix de modélisation: **(C2) la discrétisation des notes se fait en trois quantiles: Insuffisance (regroupant VI, I et FI), Normal et Excès (regroupant VE, E et FE)**.

Les données décrivent un procédé de transformation, qui s'inscrit dans le temps. Ainsi, les mesures du pétrissage sont réalisées avant les mesures de l'apprêt, et ainsi de suite. Nous posons donc une troisièmecontrainte: **(C3) seules les mesures effectuées avant une mesure donnée peuvent avoir un impact sur la valeur de celle-ci**. Ainsi, on considère que si la mesure de ***Pétrissage*** - *Consistance* peut avoir un impact sur ***Apprêt*** - *Activité fermentaire*, l'inverse n'est pas vrai. Par ce travail, nous nous plaçons déjà dans une optique de modélisation: plus que de représenter les liens (et donc les potentielles corrélations pouvant exister), nous cherchons également à déterminer les notes critiques qui pourraient avoir un impact sur les autres. Le choix de modélisation peut également être précisé: **(C3.1) une mesure A peut potentiellement expliquer une mesure B si et seulement si A et B appartiennent à une même étape, ou alors si A est réalisée dans une étape antérieure à l'étape de B.**

Connaissances Expertes Des entretiens ont été menés auprès d'experts de la profession et du domaine. Au total, cinq experts issus de différents cadres (meunerie, boulangerie, laboratoire) ont pu être consultés durant des entretiens individuels afin d'éliciter les liens et relations déjà connus entre certaines notes. Les contraintes exprimées ont ainsi pu être intégrées pour la phase d'apprentissage de structure:

- "Je sais que les mesures A et B sont indépendantes". Dans ce cas, on peut exclure tous les modèles où des liens existent entre ces variables.
- "Je sais que la mesure A peut expliquer la mesure B, mais pas le contraire". Dans ce cas, on peut exclure tous les modèles où des liens B→A apparaissent.
- "Je sais que les mesures A et B sont expliquées par une variable commune." Dans ce cas, on considère qu'elles sont corrélées, et qu'on peut donc exclure tous les modèles où des liens existent entre ces variables.

Ces connaissances sont utiles pour deux cas:

1. **Apprentissage du modèle.** Ces contraintes permettent de guider l'apprentissage de la structure et ainsi sélectionner celle s'approchant le plus de la réalité.
2. **Vérification du modèle.** Ces contraintes permettent de vérifier les liens appris dans le modèle et ainsi permettre de vérifier si certaines relations sont présentes - ou non.

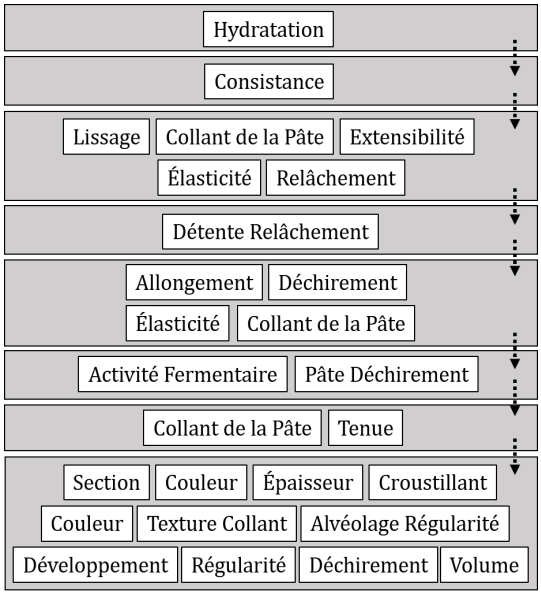

Figure 3: Schéma relationnel utilisé pour l'apprentissage du réseau Bayésien.

L'intégration des connaissances pour l'apprentissage et la vérification du modèle sera détaillé dans la section suivante.

Modélisation de l'essai de panification

Le plus gros travail de cette modélisation a concerné l'apprentissage du RB associé à l'essai de panification. Dans cette sous-section, nous présenterons la méthode appliquée, ainsi que le résultat obtenu et son évaluation. L'implémentation a été réalisée en Python, à l'aide notamment de la bibliothèque PyAgrum [5].

Apprentissage Pour apprendre le RB, le SR décrit en Figure 3 a été utilisé. Celui-ci est un mélange entre d'une part les contraintes temporelles posées par **(H3)**, mais également quelques connaissances expertes: on considère ainsi que ***Pétrissage*** - *Hydratation* a un potentiel explicatif sur l'ensemble des autres variables, mais pas l'inverse; de même, il est généralement considéré que ***Pétrissage*** - *Consistance* dépend de l'hydratation, mais peut expliquer les autres variables au **Pétrissage**. Ce choix est directement lié à la façon dont un test est réalisé: l'eau est en effet ajoutée de façon à viser une valeur de ***Pétrissage*** - *Consistance* de 10 note qui reflète un comportement normal pour la profession. Par défaut, cette dernière devrait donc être une constante: dans les faits, il arrive qu'elle ne le soit pas, marquant ainsi des blés très particuliers, à caractère problématiques pour la suite du test. Il s'agit donc d'une information cruciale pour la suite de son déroulé.

Résultat et Évaluation

Le RB appris est présenté en Figure 4. Comme indiqué en Section 3, l'apprentissage de la structure et des paramètres a été réalisé contraint par le SR, via l'algorithme Greedy

Hill Climbing guidé par un score BIC. Deux évaluations sont possibles:

1. **Évaluation qualitative.** A partir de la structure apprise, on présente le réseau à des experts afin de déterminer si les liens appris semblent cohérents avec leur connaissance des notes.
2. **Évaluation numérique.** Une validation croisée à cinq blocs est effectuée: à partir de la structure apprise avec toutes les données, nous apprenons de nouveaux paramètres sur 80% du jeu de données. Pour chaque ligne des 20% restants, nous estimons ensuite à quel point nous pouvons prédire le comportement de l'essai de panification en connaissant les notes au Pétrissage, afin d'évaluer l'efficacité du modèle décrit en Figure 1. Le test est répété cinq fois, en faisant tourner les données sélectionnées.

Évaluation qualitative. La principale difficulté de cette analyse repose sur la complexité du modèle appris: si l'évaluation de relations avec seulement un ou deux parents peut encore être possible, les variables à trois ou quatre parents sont tout de suite plus complexes à évaluer pour les experts. Dans cette partie, nous avons fait le choix de ne pas présenter les tables de probabilités conditionelles, et nous nous sommes concentrés sur la structure du modèle. Si dans le cadre de cette étude, un retour complet n'a pas encore pu être fait auprès des principaux experts (notamment ceux ayant réalisé la plupart des tests présents dans la base d'apprentissage), certains liens ont pu être vérifiés auprès de quelques autres, plus particulièrement spécialistes des premières étapes (**Pétrissage**, **Apprêt**, **Façonnage**). Afin d'aider leur analyse, un calcul d'influence a été fait pour aider la visualisation: puisque les données traitées sont ordinales (Insuffisance $\prec$ Norme $\prec$ Excès), alors il est possible dans le cas de relations monotones simples d'évaluer l'évolution de l'espérance pour chaque relation parent-enfant. Ainsi, dans la table présentée en Figure 5, on peut voir que que la valeur de ***Pétrissage***_*Relachement* augmente (passant de Normal à Excès), la valeur moyenne de ***Pétrissage***_*CollantDeLaPate* augmente également (passant de Normal à Excès). Un extrait du réseau appris coloré afin d'illustrer ces dépendances est présenté en Figure 6.

Évaluation Numérique. Afin d'évaluer le caractère prédictif du modèle, nous avons cherché à prédire chaque note en fonction des notes de pétrissage observées. La valeur modale a été choisie en post-traitement pour estimer les notes. L'évaluation du modèle est basée sur le calcul du F-score permettant de fournir un bon retour sur la fiabilité des prédictions. Ce F-score est ensuite comparé à deux autres F-scores issus du hasard:

- (1) hasard équiprobable *rd* (pour chaque prédiction, une valeur possible est tirée aléatoirement de façon équiprobable)
- (2) hasard pondéré *rdp* (qui revient à considérer toutes les variables indépendantes et à prédire la valeur des notes selon leur lois marginales)

Les résultats obtenus sont présentés en Figure 7. Les F-scores obtenus représentent la valeur moyenne des F-scores calculés sur cinq répétitions. Deux façons de calculer le F-score sont représentés:

- **Macro** permet de représenter, pour chaque note, la moyenne des F-scores obtenus pour chaque valeur possible.
- **Weighted** permet de représenter, pour chaque note, la moyenne des F-scores obtenus pour chaque valeur possible, pondérée par leur représentativité dans le jeu de données.

Ainsi, si **macro** donne une idée des performances du réseau de façon globale, **weighted** permet de prendre en compte la faible représentativité de certaines valeurs. D'une façon générale, le score **weighted** est plus élevé que le **macro**: cela peut s'expliquer par le fait que certaines catégories très peu représentées sont mal prédites. Il est interessant de noter également que le F-score correspondant aux prédictions (*pred*) est plus élevé que le score aléatoire pondéré (*rdp*), mais principalement pour les premières étapes après le **pétrissage**. On a donc bien un apport d'information nous permettant de prédire le comportement de la pâte au **pétrissage**, mais celui-ci n'est pas suffisant pour déterminer l'ensemble de l'essai de panifications: d'autres facteurs ont donc une influence qui n'a pas été capturée ici. Cela est particulièrement remarquable dans le cas de l'*activité fermentaire* à l'**apprêt** (*A_Act*): tous les scores sont assez bas, et le hasard en obtient un plus haut que le modèle. Dans ce cas-ci, le modèle n'est pas suffisant pour déterminer ce comportement. Il reste néanmoins pertinent pour informer sur les notes obtenues au **pointage** et au **façonnage**.

Discussion

Les résultats obtenus montrent bien la difficulté présente pour prédire les résultats de l'essai de panification. Malgré un jeu de données conséquent et un problème simplifié (puisque l'on s'intéresse ici à la prédiction de tendances Insuffisance/Norme/Excès et non des notes précises), de nombreux problèmes sont encore rencontrés lors de l'évaluation:

P1 Les notes du **Pétrissage** apportent des informations permettant d'estimer certaines notes obtenues dans les étapes juste après, mais pas au délà. Cela se traduit par des performances équivalentes au hasard sur certaines notes (***Façonnage***_*Déchirement*, plusieurs notes de Mie), voire inférieure dans quelques cas (***Apprêt***_*ActivitéFermentaire*, ***Mie***_*Alvéolage*).

P2 Le modèle prédit mal les cas où les notes ne sont pas "normales": si l'on ne regarde que les notes différentes de Normal, celles-ci sont généralement mal prédites, avec des performances parfois moins bonnes que le hasard. Cela est dû à deux choses: d'une part, la valeur modale favorise généralement les valeurs "normales"; d'autre part, certains cas sont proches de l'équiprobabilité, et difficiles à trancher.

P3 L'équiprobabilité de certains cas montre bien les limites de la prédictibilité du comportement de

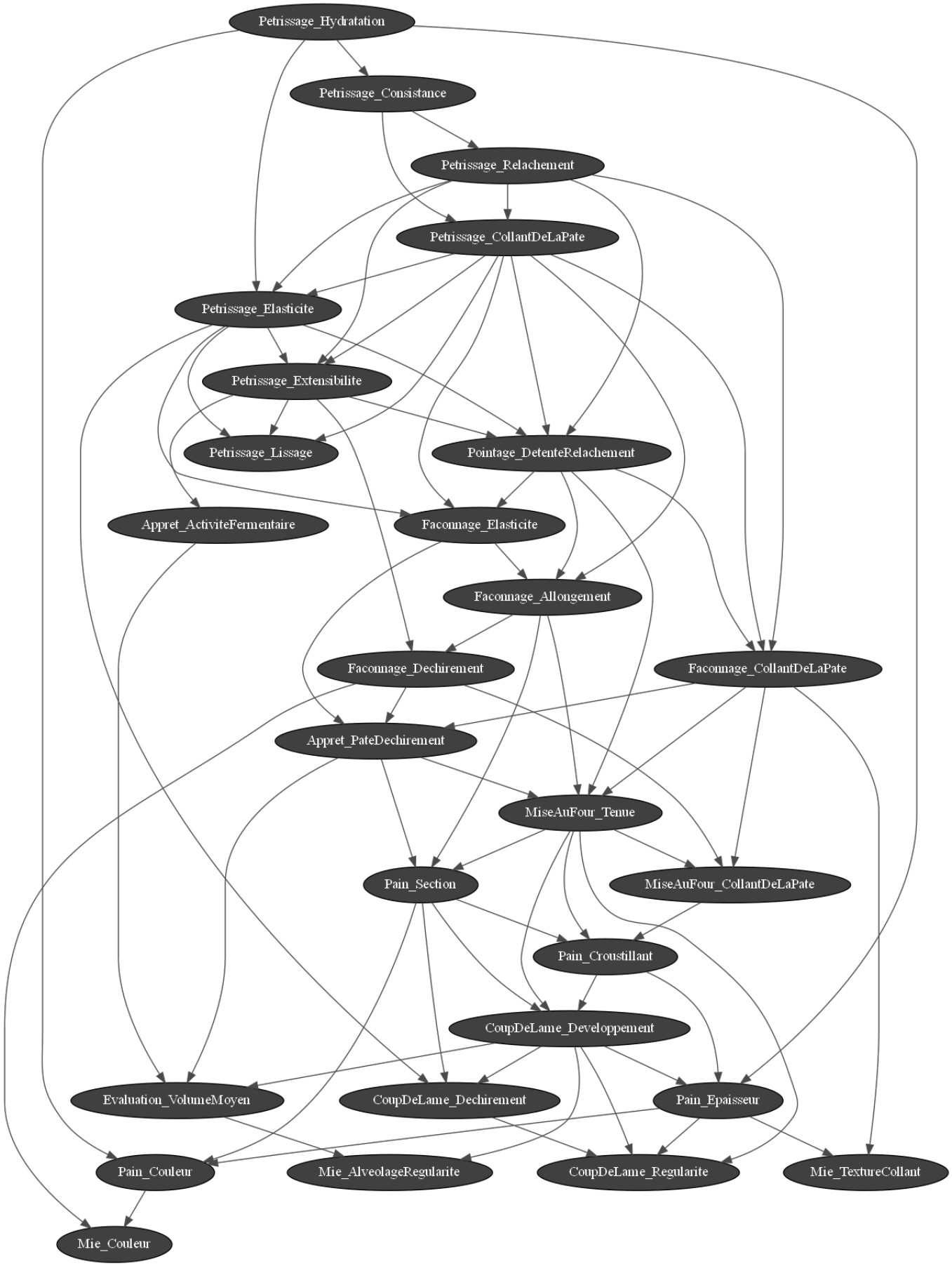

Figure 4: RB appris a partir des données et des contraintes expertes.

		Petrissage_CollantDeLaPate	
Petrissage_Relachement	Petrissage_Consistance	aN	bFE_E_VE
aN	aVI_I_FI	0.5912	0.4088
	bN	0.7691	0.2309
	cFE_E	0.9435	0.0565
bFE_E_VE	aVI_I_FI	0.1193	0.8807
	bN	0.0787	0.9213
	cFE_E	0.2857	0.7143

Figure 5: Table de probabilités conditionnelles pour la variable ***Petrissage***_*CollantDeLaPate*. Les noms des valeurs sont construits telles que [indicatif d'ordre][ensemble des valeurs prises].

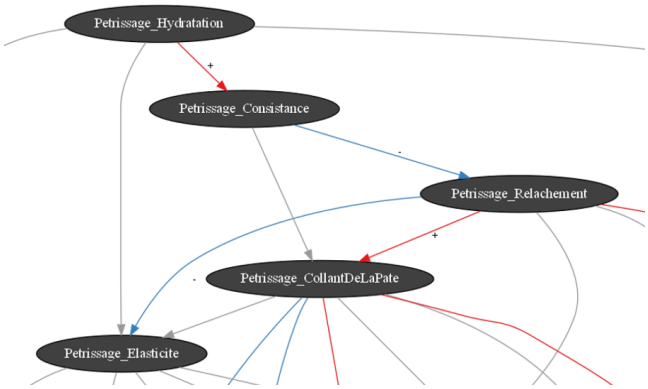

Figure 6: Exemple de relations monotones positives (rouge) et négatives (bleu) entre variables. Cette visualisation a été proposée aux experts afin de faciliter leurs retours.

la pâte lors de l'essai de panification. Ainsi, la variable ***Façonnage***_*Déchirement* dépend de deux autres variables: ***Façonnage***_*Allongement* et ***Pétrissage***_*Extensibilité*. Néanmoins, comme montré en Figure 8, plusieurs combinaisons de valeurs pour ces parents résultent en des probabilités très proches, et calculées sur un échantillon conséquent (ainsi, lorsque tous ses parents sont excessifs, on dénombre 400 échantillons où ***Façonnage***_*Déchirement* est en excès, et 393 normaux).

Afin de développer ce modèles, plusieurs solutions sont donc à explorer:

S1 Des informations supplémentaires doivent être apportées au réseau afin de mieux prédire les notes trop éloignées du **pétrissage**. Cela pourrait se faire par l'étude de la prédictabilité de certaines notes via d'autres critères descriptifs du blé: à la manière du travail effectué par ASCOPAIN, il serait intéressant de pouvoir avoir une idée des valeurs prises par d'autres notes afin de renforcer les informations apportées dans le RB.

S2 L'objectif de ce modèle étant de prévenir les soucis pouvant être rencontrés lors de l'essai de panification, il serait intéressant de concentrer l'apprentissage sur les cas problématiques, en diminuant la représentativité des cas "normaux". Un travail a été initié auprès d'experts afin de déterminer les profils complets jugés non problématiques: il pourrait être pertinent de concentrer l'apprentissage du modèle sur les autres profils.

S3 Tout comme pour S1, l'ajout d'informations supplémentaires aiderait à trancher sur les cas équiprobaux. Il serait également intéressant de présenter ces cas particuliers aux experts, afin de comprendre leur origine et éventuellement suggérer des expériences, critères, ... permettant de mieux comprendre le comportement de la pâte à ces moments.

S4 Comme évoqué plus haut, la validation experte peut se heurter facilement à la ocmplexité du modèle appris. Le développement d'outils de visualisation - tels que la coloration des relations monotones - est une piste à explorer afin de faciliter le travail de l'expert et, par extension, améliorer les modèles appris.

S5 La technique de propagation mise en place ici repose uniquement sur les notes observées durant le Pétrissage: il est donc normal de considérer que les incertitudes s'accumulant rendent la prédiction plus difficile pour les notes plus éloignées dans le test. Une autre simulation a essayer serait alors de fixer chronologiquement les notes prédites par leur valeur modale, afin de mieux propager l'information.

5 Conclusion

Nous venons de présenter un modèle visant à prédire la qualité du blé à partir de ses critères descriptifs. Pour cela, nous avons basé notre étude sur l'essai de panification, test couramment utilisé en industrie boulangère afin de déterminer les capacités techniques d'un blé ou d'une

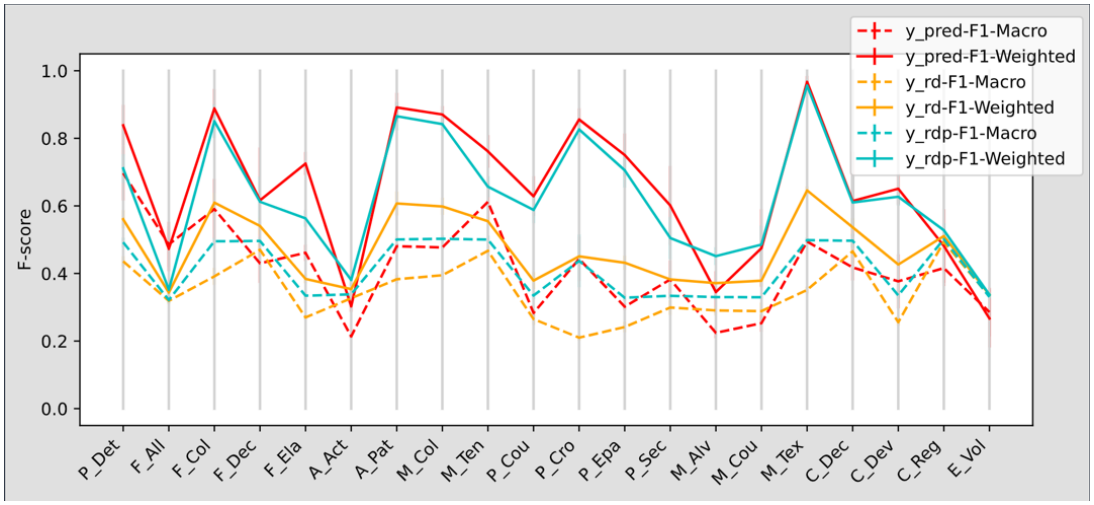

Figure 7: F-score obtenus pour chaque note. Les prédictions ont été faites sur 20% du jeu de données, à partir d'un modèle appris à partir des 80% restants. Les valeurs représentent la moyenne des F-scores obtenus via une validation croisée à cinq blocs.

		Faconnage_Dechirement	
Faconnage_Allongement	Petrissage_Extensibilite	aN	bFE_E_VE
aVI_I_FI	aVI_I_FI	0.1519	0.8481
	bN	0.2462	0.7538
	cFE_E_VE	0.2759	0.7241
bN	aVI_I_FI	0.1648	0.8352
	bN	0.3743	0.6257
	cFE_E_VE	0.3237	0.6763
cFE_E_VE	aVI_I_FI	0.3392	0.6608
	bN	0.5417	0.4583
	cFE_E_VE	0.4956	0.5044

Figure 8: Table de probabilités conditionnelles pour le Déchirement au Façonnage. Les deux dernières lignes dénotent de cas equiprobables où le réseau est incapable de trancher sur la présence ou non d'un défaut.

farine. A partir d'une base de données représentant des essais effectués sur les 20 dernières années et guidés par des connaissances expertes, nous avons pu apprendre un réseau Bayésien visant à lier les différentes notes de l'essai de panification les unes aux autres. L'objectif final étant de coupler ce modèle avec un autre pré-existant, permettant de lier les caractéristiques d'un blé à son comportement au pétrissage, nous avons alors évalué la capacité du réseau appris à prédire les notes d'un essai à partir d'un profil de pétrissage connu.

Les résultats obtenus ont montré que s'il était possible d'évaluer certaines notes à partir du profil de pétrissage, il restait des facteurs encore non pris en compte permettant de définir avec précision les notes manquantes.

De prochains travaux viseront à développer à la fois les connaissances expertes à intégrer au modèle, à la fois pour améliorer ses prédictions (en apportant par exemple des informations supplémentaires), mais également pour faciliter les retours et l'évaluation par les experts du domaine.

References

[1] Chickering, D. M. 2003. Optimal Structure Identification with Greedy Search. *J. Mach. Learn. Res.*, 3(null): 507–554.

[2] Cooper, G. F.; and Herskovits, E. 1992. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9(4): 309–347.

[3] De Campos, C.; Zhi, Z.; and Ji, Q. 2009. Structure Learning of Bayesian Networks Using Constraints. In *ICML*, 113–120.

[4] De Campos, C. P.; and Ji, Q. 2008. Improving Bayesian Network parameter learning using constraints. In *ICPR*, 1–4.

[5] Ducamp, G.; Gonzales, C.; and Wuillemin, P.-H. 2020. aGrUM/pyAgrum : a Toolbox to Build Models and Algorithms for Probabilistic Graphical Models in Python. In *10th International Conference on Probabilistic Graphical Models*, volume 138 of *Proceedings of Machine Learning Research*, 609–612. Skørping, Denmark.

[6] Friedman, N.; Getoor, L.; Koller, D.; and Pfeffer, A. 1999. Learning Probabilistic Relational Models. In *IJCAI*, 1300–1307. Morgan Kaufmann Publishers Inc.

[7] Kansou, K.; Chiron, H.; Della Valle, G.; Ndiaye, A.; and Roussel, P. 2014. Predicting the quality of wheat flour dough at mixing using an expert system. *Food Research International*, 64: 772–782.

[8] Munch, M.; Dibie, J.; Wuillemin, P.; and Manfredotti, C. E. 2019. Towards Interactive Causal Relation Discovery Driven by an Ontology. In *FLAIRS*, 504–508.

[9] Munch, M.; Wuillemin, P.-H.; Manfredotti, C.; Dibie, J.; and Dervaux, S. 2017. Learning Probabilistic Relational Models Using an Ontology of Transformation

Processes. In *OTM 2017 Conferences*, 198–215. ISBN 978-3-319-69459-7.

[10] Parviainen, P.; and Koivisto, M. 2013. Finding Optimal Bayesian Networks Using Precedence Constraints. *Journal of Machine Learning Research*, 14: 1387–1415.

[11] Schwarz, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2): 461 – 464.