

Association Suisse des Actuaires Assurances du personnel



Construction de tables de turnover par application de l'approche d'apprentissage automatique dans l'évaluation des IFC en norme IAS 19 *

Minh Tu Pham

Consultant senior en actuariat – Wealth Solutions – Aon, Paris

Actuaire qualifié de l'Institut des Actuaires (IA), France

() Prix : Un des meilleurs mémoires d'actuaire de l'année 2021 de l'IA*

Andermatt, 26 Août 2022

- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Référentiel

- Une norme internationale dédiée à la valorisation des passifs des entreprises relatifs aux avantages du personnel ;
- Définir les méthodes d'évaluation et de comptabilisation à appliquer. (cf. Annexes)

Champs d'application

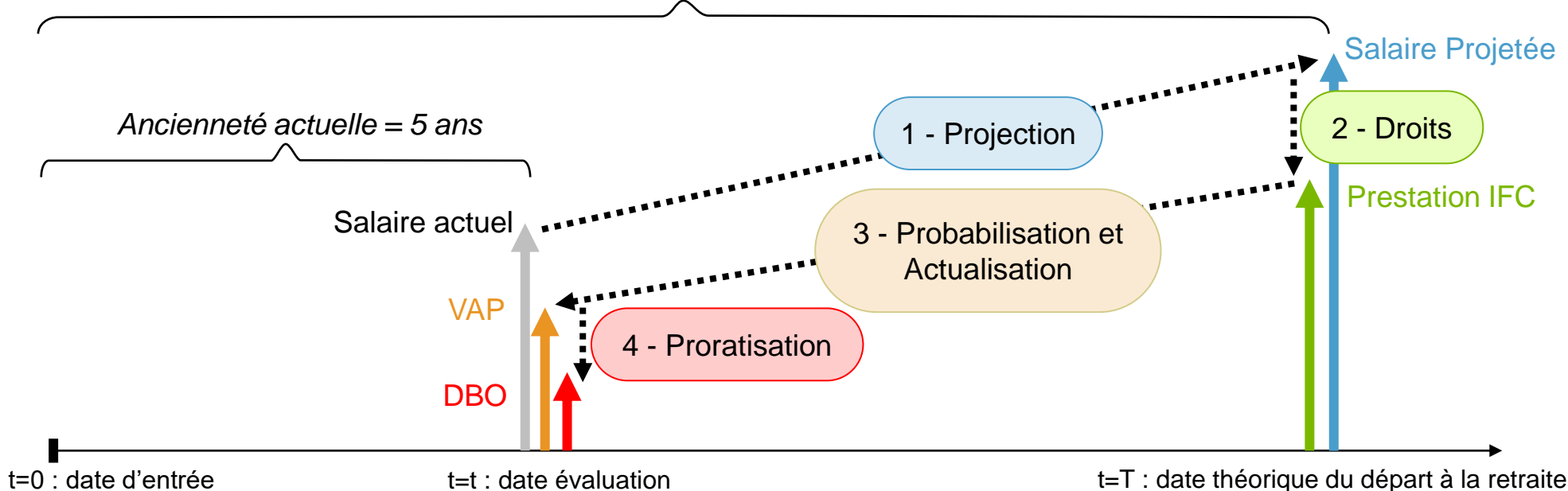
- Les avantages à court terme ;
- Les avantages postérieurs à l'emploi ;
- Les autres avantages à long terme ;
- Les indemnités de cessation d'emploi.

Indemnités de Fin de Carrière (IFC)

- L'IFC est un régime à **prestations définies** qui garantit au salarié le versement d'une **indemnité sous forme de capital** lors de son départ à la retraite.
- Méthode de calcul de l'engagement des IFC : **Méthode PUC(*) avec services proratisés**

$$\text{Engagement}(DBO) = VAP * \frac{\text{Ancienneté actuelle}}{\text{Ancienneté à la retraite}}$$

Ancienneté à la retraite = 20 ans



Note : la hauteur des flèches est purement représentative.

(*) Projected Unit Credit (Unité de Crédit projetée). (Attention avec la clarification de l'IFRIC en 2021)

- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Contexte et motivations de l'étude

- L'hypothèse de turnover vise à prendre en compte dans l'engagement la **probabilité** que le salarié **soit encore présent** dans l'entreprise au moment de son départ à la retraite.

Pratiques marché et limites

- ? Comment elle est construite en pratique ? Estimer le **taux de sortie à chaque âge** pour une population concernée
- ☹ De **nombreux écarts** entre le turnover choisi dans l'hypothèse et le turnover réel
- + **Modélisation de turnover** : modèles, facteurs importants non pris en compte

Exigences de la norme IAS 19 et des CAC en France

- ❖ L'hypothèse **n'est pas clairement définie** dans la norme
- ❖ La note de la CNCC : **seuls les motifs de démissions** doivent être pris en compte (cf. Annexes)

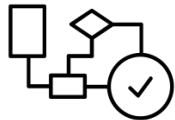
Quelques avantages

- ✓ Développements de techniques de **machine learning**
- ✓ Puissance de calculs des ordinateurs
- ✓ Volume vaste des données disponibles

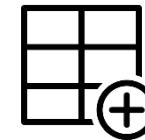
Nouvelle approche pour déterminer les taux de turnover **au niveau individuel** avec l'aide des modèles d'apprentissage automatique.

Objectifs

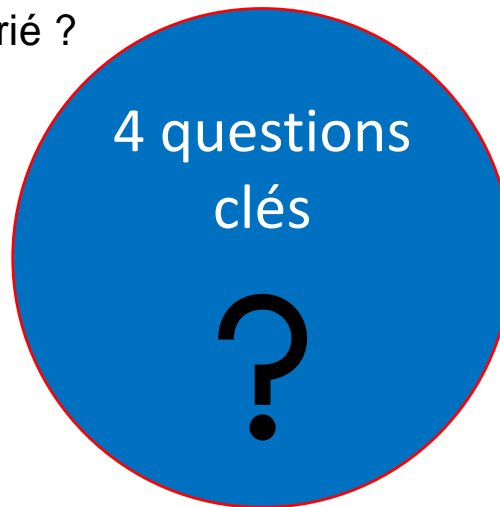
- Examiner les **opportunités** des techniques d'apprentissage automatique actuelles ;
- Voir comment ces techniques peuvent **aider à construire des tables de turnover** qui **reflètent les meilleures estimations** de rotation du personnel au sein de l'entreprise.



Algorithme le plus approprié ?



Comment construire de nouvelles tables de turnover ?



Prédicteurs importants ?



Impacts sur l'évaluation des IFC ?

Modèle de référence de l'évaluation des IFC (1/2)

- Les hypothèses actuarielles retenues :

Hypothèses	2019
Date de calcul	31/12/2019
Hypothèses financières	
Taux d'actualisation	0,75%
Inflation	1,50%
Taux de revalorisation des salaires (inflation incluse)	Cadres : 2,50% Non cadres : 2,00%
Taux de charges sociales	49,50%
Hypothèses démographiques	
Tables de mortalité	INSEE H/F 2014-2016
Turnover	Taux en fonction de l'âge, la CSP et l'entité (cf. Annexes)
Âge de départ à la retraite	Age taux plein
Modalité de départ	Départ volontaire
Droits évalués	En mois de salaire par ancienneté

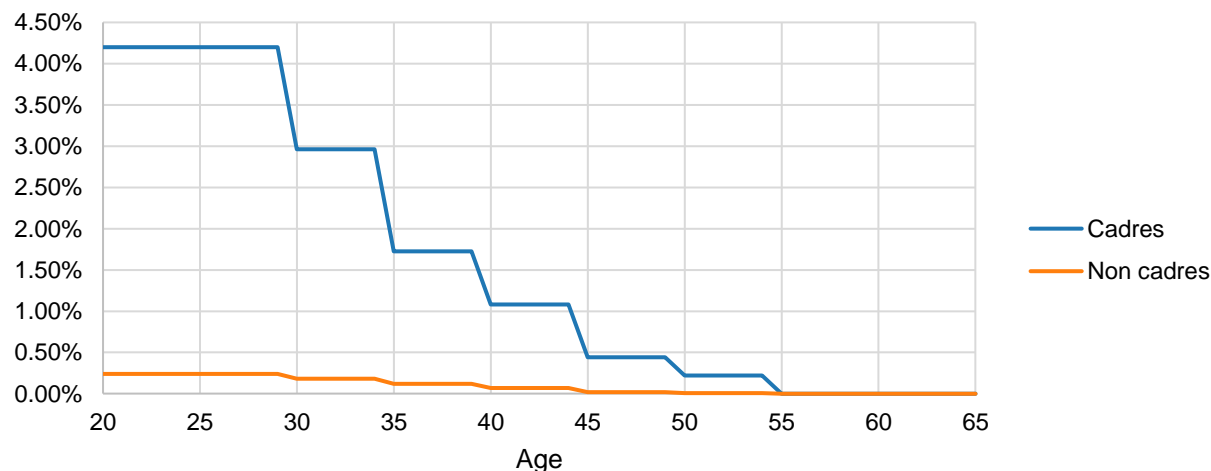


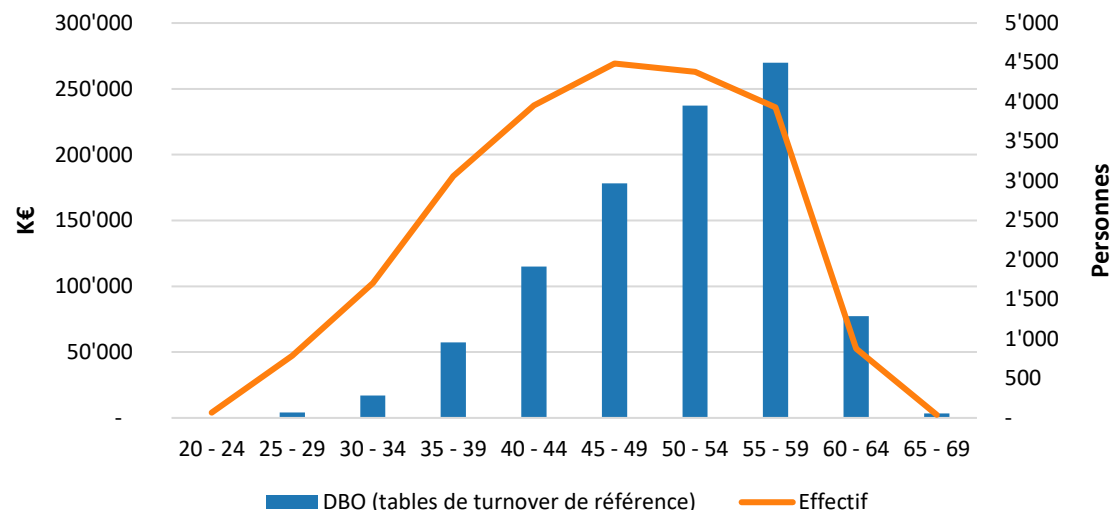
Figure : Taux de turnover de l'entité S126

Modèle de référence de l'évaluation des IFC (2/2)

- Les résultats actuariels (*montants en milliers d'euros*) :

Catégorie	Effectif	Engagement (DBO) au 31/12/2019	Coût de services (SC) 2020	Coût d'intérêt (IC) 2020	Charge 2020 (SC + IC)
Cadres	9 837	567 510	34 746	4 223	38 969
Non cadres	13 461	392 196	21 235	2 890	24 125
Total	23 298	959 706	55 981	7 113	63 094

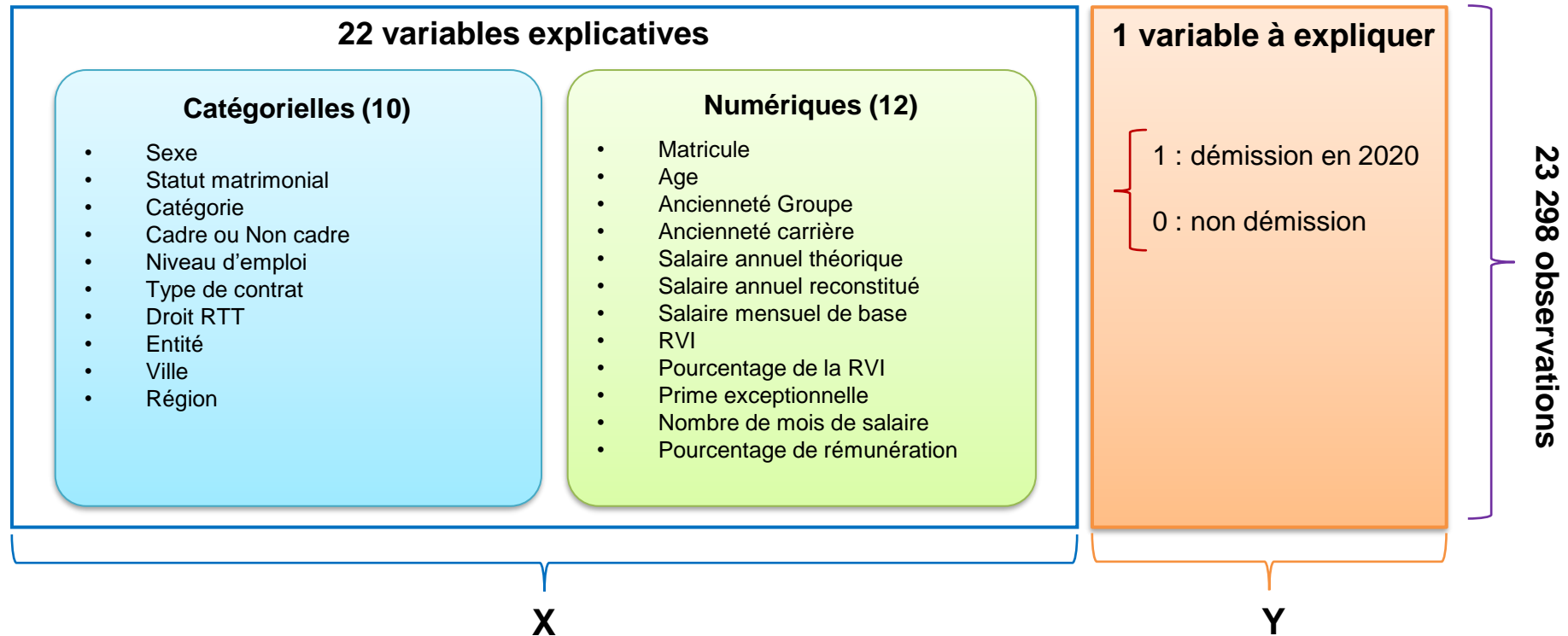
- Répartition de la population et de l'engagement au 31/12/2019 par tranche d'âge :



- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Analyses exploratoires de données



- Contrôles sur les données
- Analyses statistiques de chaque variable
- Etudes de corrélation entre les variables

→ **Comprendre** les données et **Anticiper** des problèmes potentiels

→ **Proposer** des traitements de données et des modèles appropriés.

Variable à expliquer - Turnover

- Seulement **4,13 %**, soit 963 sur 23 298 salariés, ont été identifiés comme sortants avec pour motif « **démission** ».

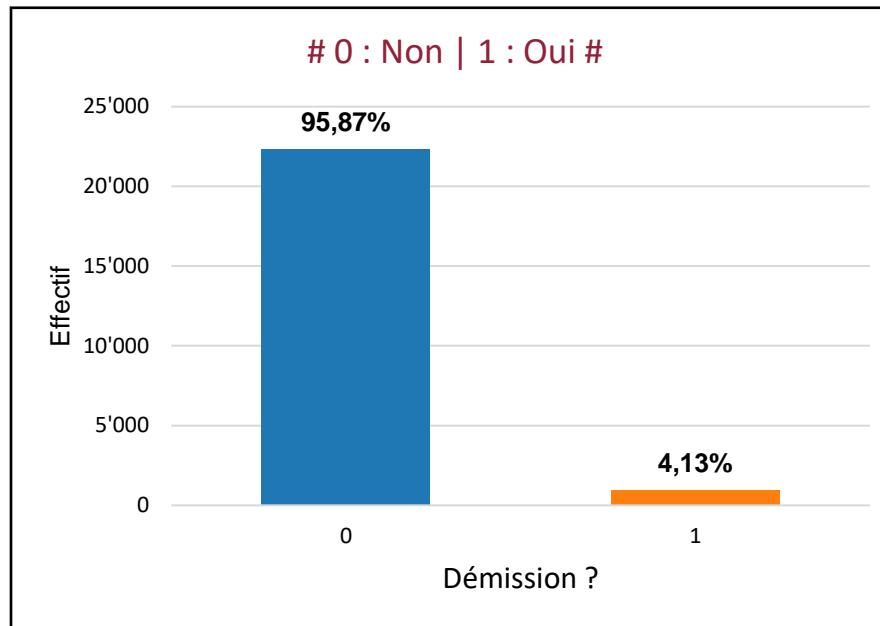
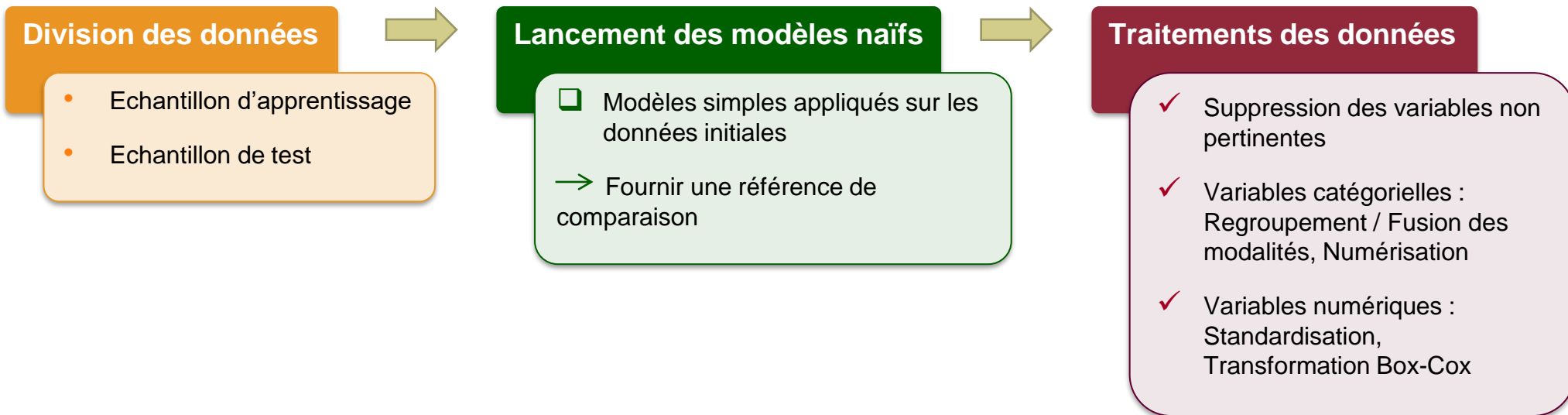


Figure : Distribution de démission

- Les modèles d'apprentissage automatique appropriés : **classification**
- Besoin de techniques afin de traiter le **problème de classification déséquilibrée**.

Traitements de données

- Les étapes principales du processus de traitements des données :



→ Améliorer la performance des modèles d'apprentissage automatique pour fournir de meilleurs résultats de prévision.

On définit un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. L'objectif est d'estimer une fonction $f: \mathbb{R}^p \rightarrow [0, 1]$ telle que f désigne la probabilité d'une observation labelisée 1.

$$\mathbb{P}(Y = 1|X) = f(X)$$

Où :

- X : Variable vectorielle $\Omega \rightarrow \mathbb{R}^p$
- Y : Variable à expliquer $\Omega \rightarrow \{0, 1\}$
- $\mathcal{D}_n = (\mathbf{X}, \mathbf{y}) = \{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$: Un ensemble de données i.i.d.

Modèles d'apprentissage automatique

- Les modèles utilisés dans cette étude :

Classification linéaire

- Régression logistique
- Lasso
- Ridge
- Elastic net

Classification non-linéaire

- K plus proches voisins
- Machines à vecteurs de support
- Réseau de neurones

CART et Modèles d'ensemble

- CART
- Forêt aléatoire
- Extreme Gradient Boosting

- Techniques pour traiter les données déséquilibrées : **Rééchantillonnage** (Sous / Sur) et **Balanced Bagging** (cf. Annexes).
- Agrégation de modèles : Généralisation empilée - **Stacking model** (cf. Annexes).

- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Comparaison des modèles d'apprentissage automatique

	Logistique	Lasso	Ridge	Elastic net	KNN	SVC	Réseau de neurones	CART	Forêt aléatoire	XGB
Echantillon initial (Modèles naïfs)	0,8285				0,6009			0,65922		
Echantillon initial traité	0,8387	0,8385	0,8382	0,8357	0,6828	0,7064	0,8479	0,6101	0,8515	0,8585
Sur-échantillonnage	0,8391	0,8408	0,8409	0,7553	0,6849	0,8192	0,7945	0,5777	0,8458	0,8443
Sous-échantillonnage	0,8346	0,8401	0,8402	0,7267	0,7762	0,7904	0,7975	0,6911	0,8469	0,8369
Balanced (Bagging)	0,8403	0,8425	0,8423	0,8181	0,8112	0,7930	0,8467	0,8620	0,8640	0,8329

Table : AUC (sur l'échantillon de test) des divers modèles d'apprentissage combinés avec les techniques de rééchantillonnage

- Critère d'évaluation : AUC (aire sous la courbe ROC).
- Les performances des modèles d'apprentissage automatique sont meilleures que celles des modèles naïfs → l'efficacité des traitements des données.
- Les techniques de rééchantillonnage ont amélioré la qualité de prévision dans plusieurs modèles.
- Les trois modèles les plus performants : XGB, Balanced Bagging CART et Balanced Bagging Forêt aléatoire.

Calibration du modèle d'apprentissage automatique final

- Ces 3 modèles ont été agrégés par le modèle *Stacking* dont les résultats d'évaluation :

	XGB	Balanced Bagging CART	Balanced Bagging Forêt aléatoire	Stacking
AUC	0,8585	0,8620	0,8640	0,8678

Table : AUC (sur l'échantillon de test) des modèles de base et du modèle *Stacking*

- Modèle final de prévision de turnover : *Stacking*.
- Prédicteurs importants : les éléments de salaire et de primes, l'ancienneté groupe, l'âge, le statut matrimonial et le sexe.

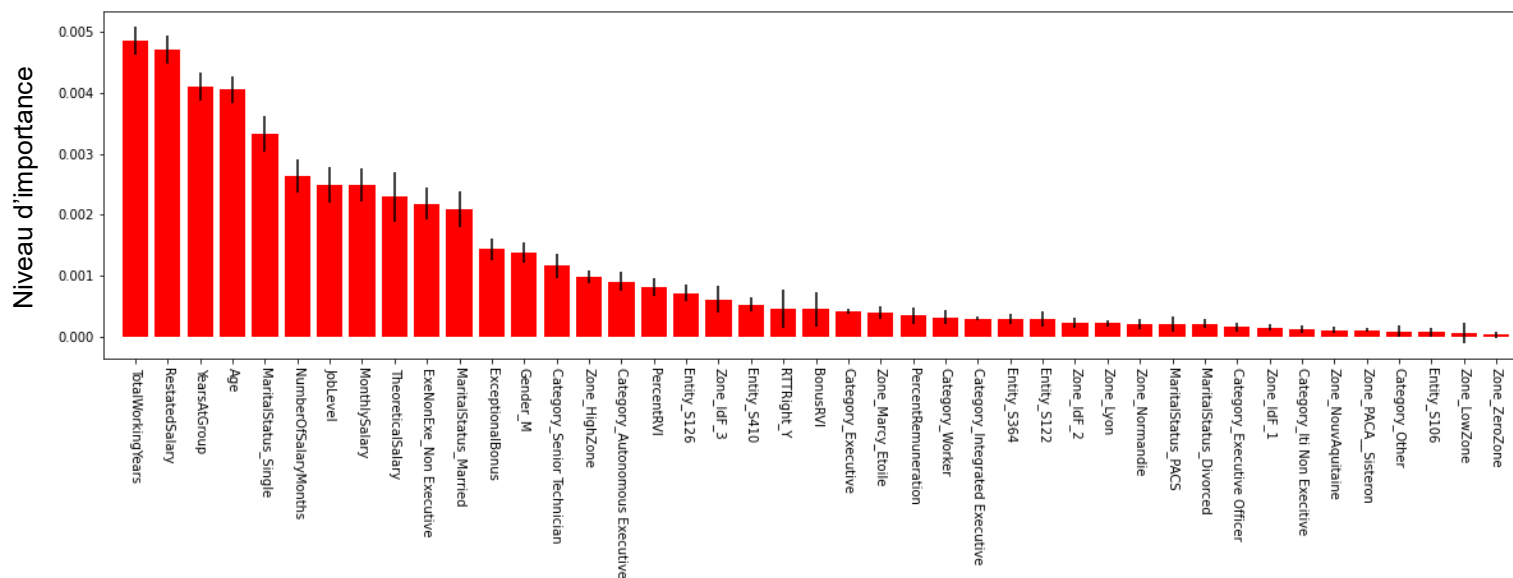


Figure : Les plus importants prédicteurs détectés par le modèle *Stacking*

Construction de nouvelles tables de turnover

- En combinant le modèle de *Stacking* et les hypothèses de projection, les 2 types de nouvelles tables de turnover ont été construits :

Les courbes de taux de turnover individuelles

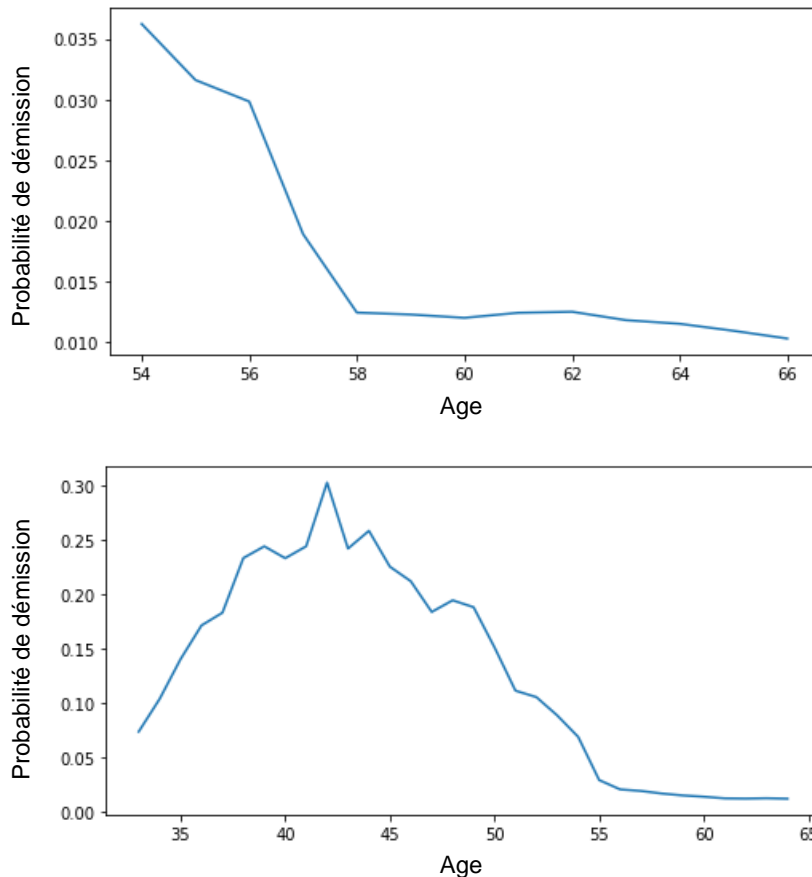


Figure : 2 exemples des courbes de taux de turnover individuelles

Une table de turnover prospective en agrégeant les courbes individuelles

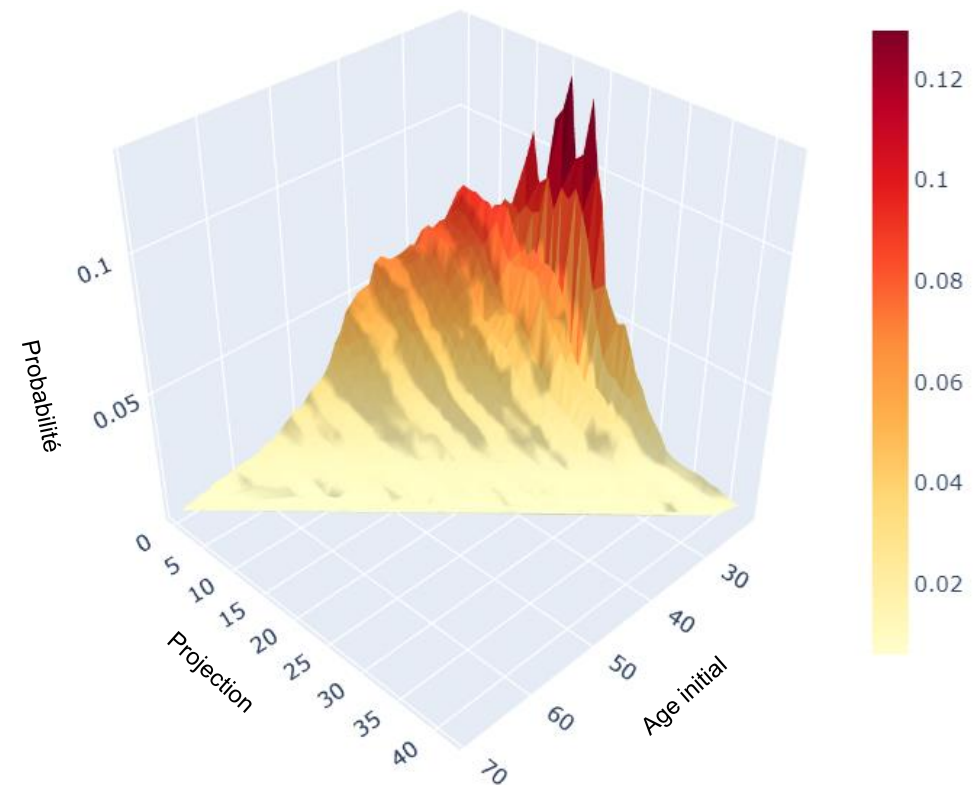


Figure : Taux de turnover prospective

Impacts des nouvelles tables de turnover sur l'évaluation des IFC

Modèles	Engagement (DBO) au 31/12/2019	Coût de services (SC) 2020	Coût d'intérêt (IC) 2020	Charge 2020 (SC + IC)
Modèle de référence	959 706	55 981	7 113	63 094
Modèle 1 (avec les courbes individuelles)	821 440	43 818	6 077	49 984
Ecart en K€	-138 307	-12 163	-1 036	-13 199
Ecart en %	-14%	-22%	-15%	-21%
Modèle 2 (avec la table prospective)	799 709	43 181	5 913	49 095
Ecart en K€	-159 997	-12 799	-1 200	-13 999
Ecart en %	-17%	-23%	-17%	-22%

Table : Résultats avec l'application des nouvelles tables de turnover (montants en milliers d'euros)

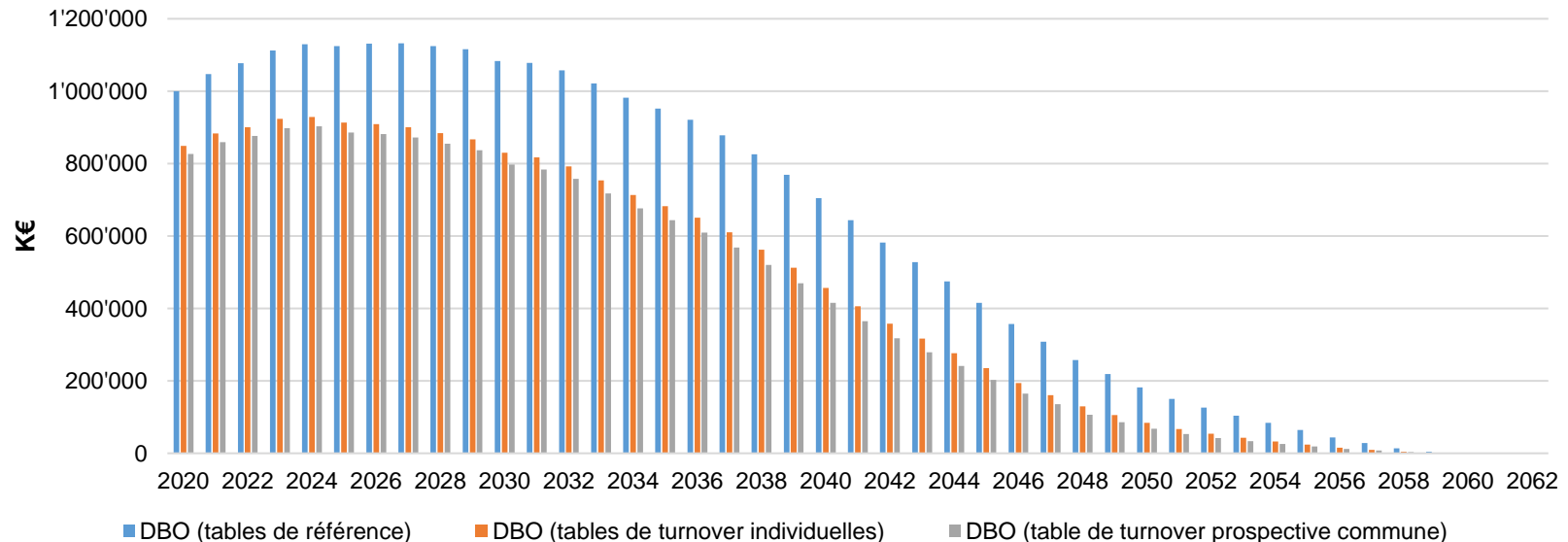


Figure : Comparaison des engagements projetés

- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Conclusions

- Une nouvelle approche pour déterminer **les taux de turnover au niveau individuel** à l'aide des modèles d'apprentissage automatique ;
- **L'utilisation des techniques modernes et puissantes** qui permettent, dans **certaines limites (cf. Annexes)**, de bonnes estimations du taux de démission ;
- **Plusieurs facteurs** autres que l'âge et la catégorie ont un **impact significatif sur la décision de démission** d'un salarié ;
- **L'hypothèse de turnover** présente un **impact important sur l'engagement** et donc sur la provision des passifs sociaux de l'entreprise ;
- **Cette hypothèse doit être revue régulièrement** par les actuaires de manière à refléter l'expérience de l'entreprise.

Merci de votre attention !

- 1 Introduction du cadre de l'étude
- 2 Motivations de l'étude et Modèle de référence
- 3 Ingénierie de données et Modèles d'apprentissage automatique
- 4 Résultats numériques et Application sur les calculs actuariels
- 5 Conclusions

Annexes

Méthodes d'évaluation des IFC en norme IAS 19

- La valeur actuelle probable de l'engagement de l'IFC de l'entreprise envers un salarié à l'âge x à la fin de l'année n :

$$VAP = \frac{IFC * {}_k p_x * {}_k r_x}{(1 + i)^k}$$

- L'engagement (DBO) à la fin de l'année n , selon la méthode PUC avec services proratisés :

$$DBO_n = VAP * \frac{\text{Ancienneté actuelle}}{\text{Ancienneté à la retraite}}$$

- Le coût des services rendus :

$$SC_{n+1} = \frac{VAP}{\text{Ancienneté à la retraite}} * (1 + i)$$

- Le coût d'intérêt :

$$IC_{n+1} = \left(DBO_n - 0,5 * EBP_{n+1} \right) * i$$

Où :

k : âge retraite – âge actuel

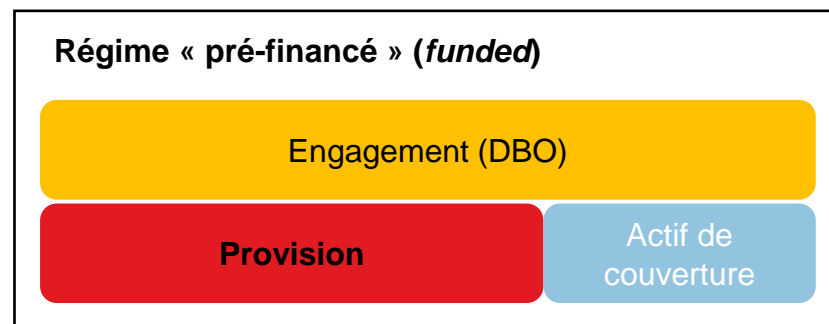
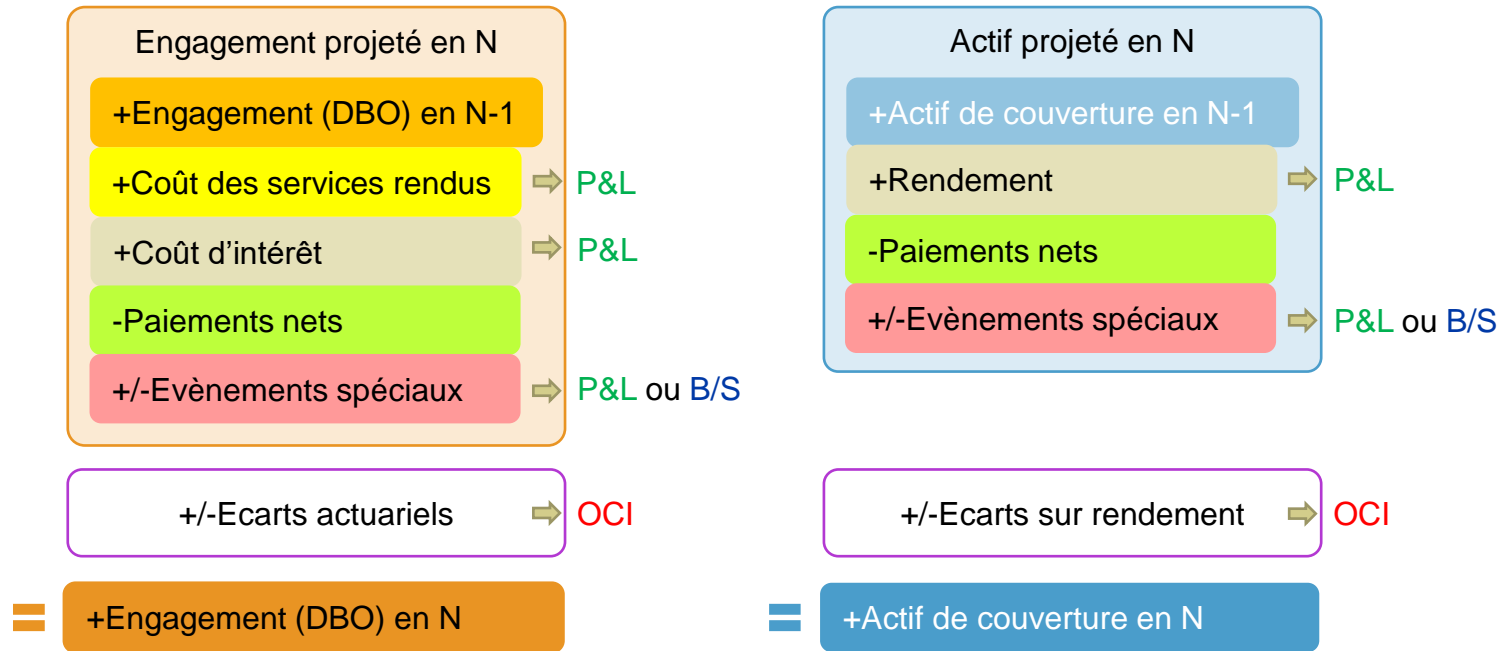
${}_k p_x$: probabilité de survie

${}_k r_x$: probabilité de présence dans l'entreprise

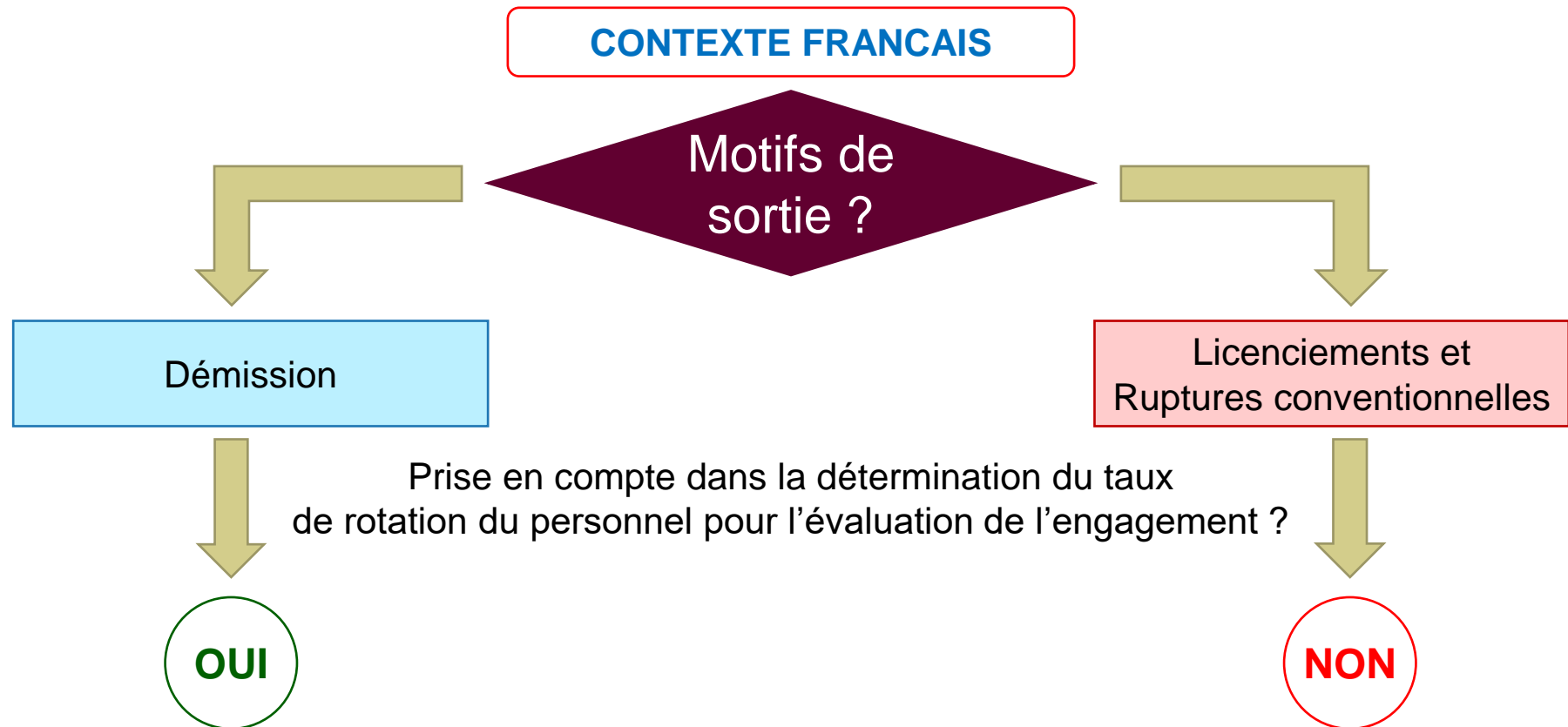
i : taux d'actualisation

EBP_{n+1} : prestation attendue de l'année $n+1$

Méthodes de comptabilisation en norme IAS 19



La note de la CNCC – Quels motifs de sortie dans le turnover ?



- *Référentiel comptable français* : Tout autre cas de départ avant l'âge de la retraite engendre, pour l'entreprise, **un paiement au moins aussi important** que l'indemnité de fin de carrière.
- *IAS 19* : Tenir compte des futurs licenciements ou ruptures conventionnelles dans le calcul du taux de rotation retenu pour calculer l'engagement de retraite aboutirait à **sous-évaluer les provisions** reconnues au bilan au titre des indemnités de départ à la retraite.

Techniques de traitement des données déséquilibrées

- Sous et sur - échantillonnage :

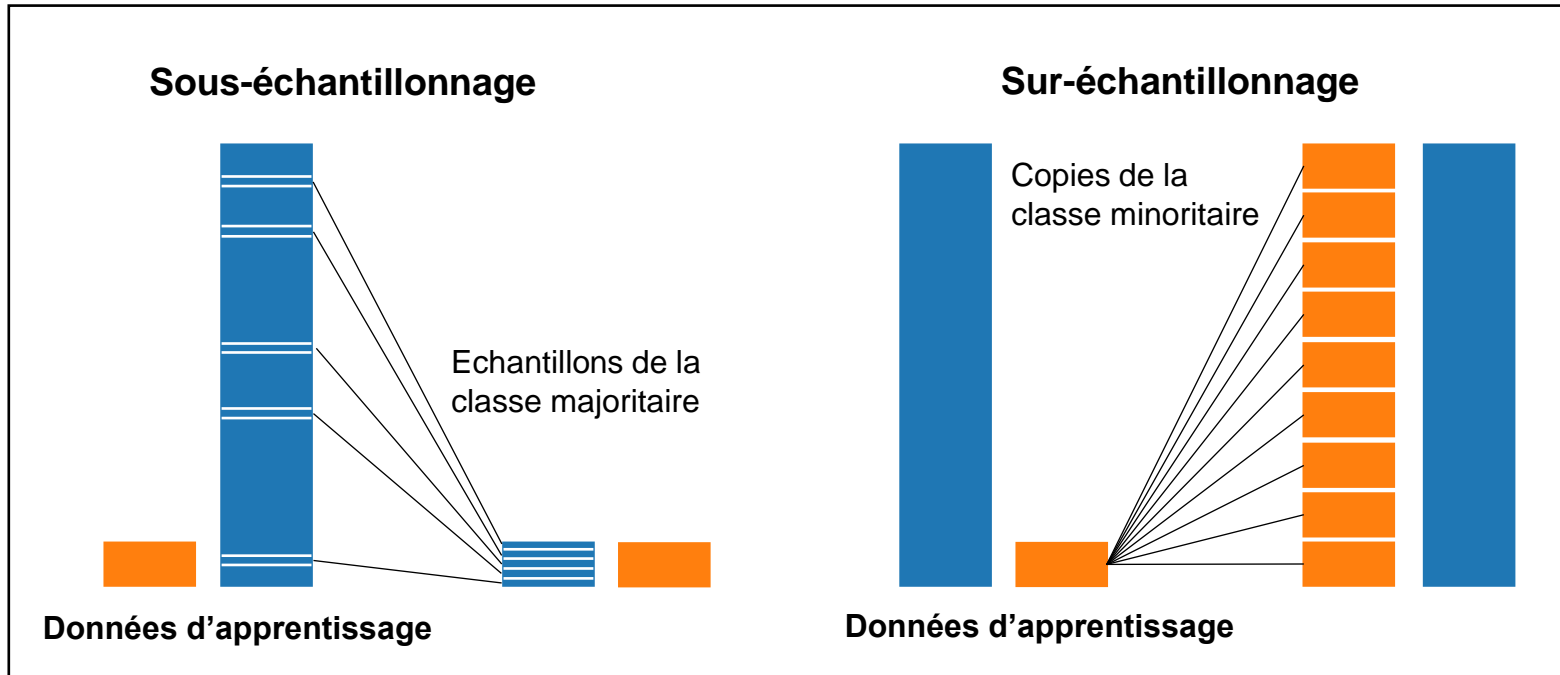


Figure : Techniques sous et sur-échantillonnage

- Techniques *Balanced Bagging* : le sous-échantillonnage est effectué à chaque étape d'échantillonnage (*bootstrapping*).

Généralisation empilée – *Stacking model*

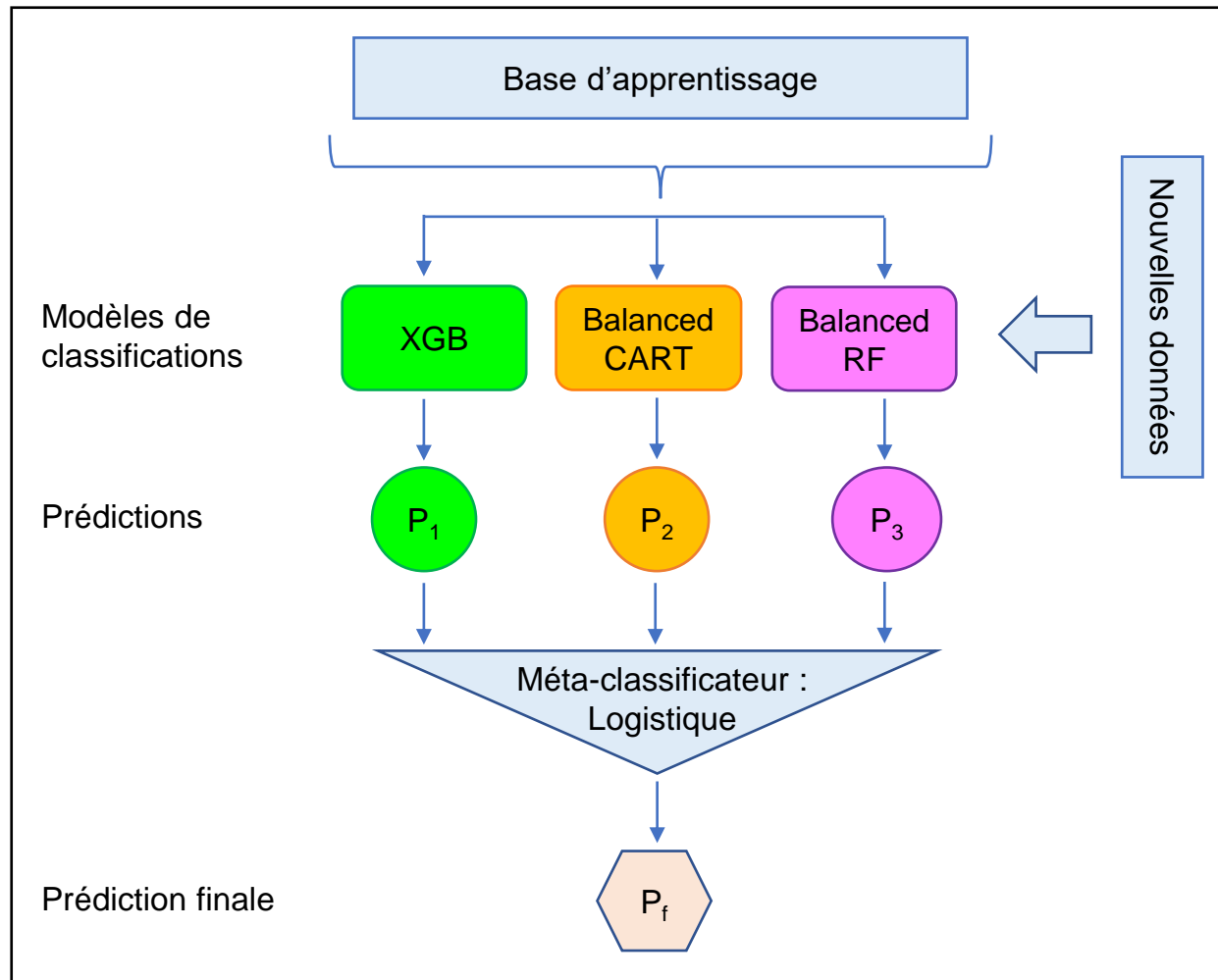
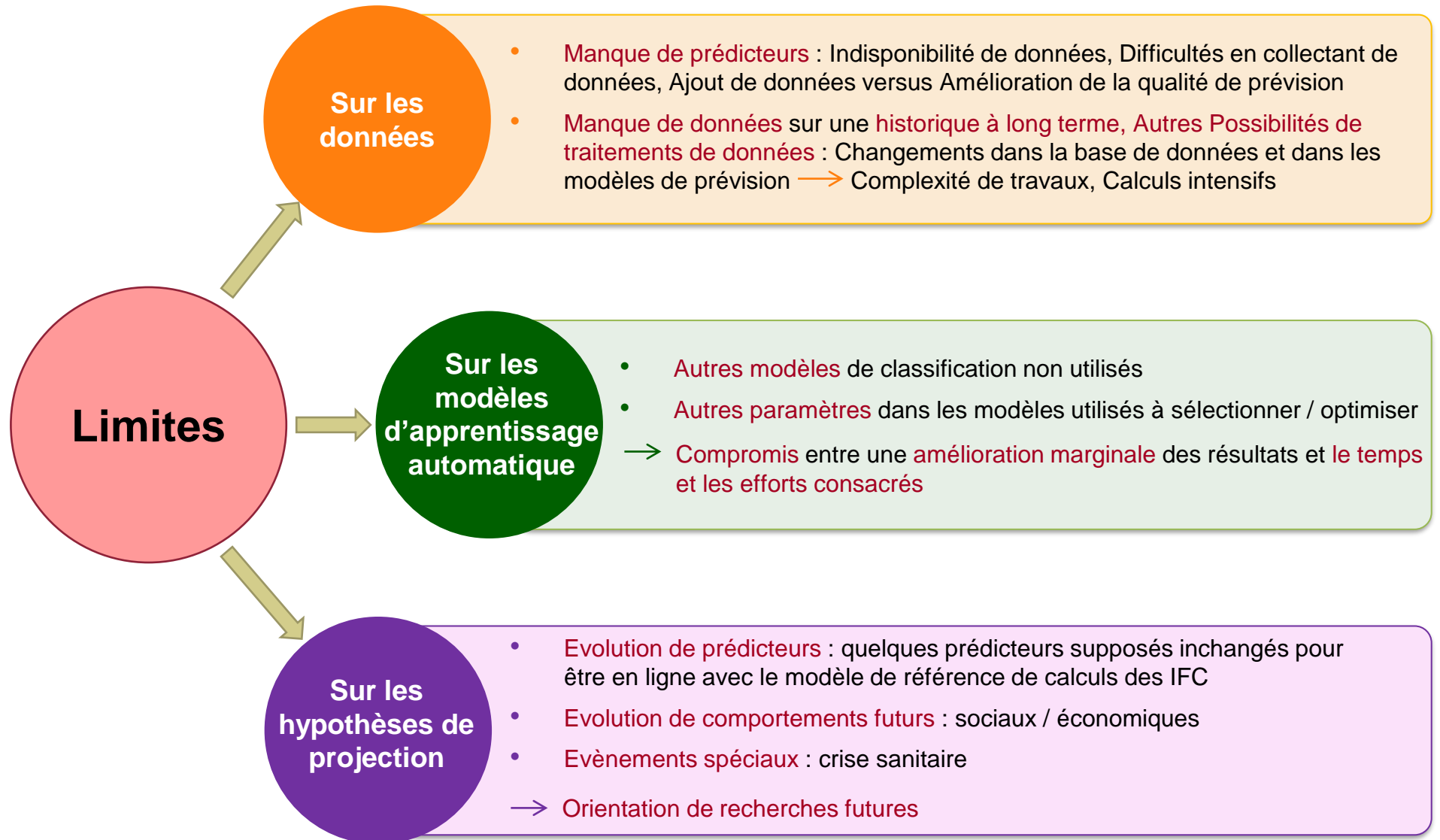


Figure : Illustration graphique de la généralisation empilée – *Stacking model*



Minh Tu PHAM

Actuaire qualifié IA | Consultant senior en actuariat

Wealth Solutions – Aon

31-35 rue de la Fédération, 75717 Paris Cedex 15, France

Tél : +33 (0)1 47 83 13 49 – Mob : +33 (0)6 26 49 93 38

E-mails : minh.tu.pham@aon.com ; phamminhtux4@gmail.com

LinkedIn : <https://www.linkedin.com/in/minh-tu-pham-123834147/>

