Évaluation de l'uplift multi-traitement sur des données biaisées dans le cas du non-random assignment

Nathan Le Boudec*,**, Nicolas Voisine*
Bruno Crémilleux**

*Orange Labs 22300 Lannion, France
{nathan.leboudec, nicolas.voisine}@orange.com,
** UNICAEN, ENSICAEN, CNRS - UMR GREYC,
Normandie Univ 14000 Caen, France
bruno.cremilleux@unicaen.fr

Résumé. L'uplift quantifie l'impact d'une action (marketing, traitement médical) sur le comportement d'un individu. La prédiction de l'uplift repose sur l'hypothèse que les groupes ciblés et le groupe de contrôle sont équivalents. Cependant, des clients sont susceptibles d'être ciblés en raison de leur comportement passé, ce qui introduit un biais et fausse l'estimation de l'uplift. Ce problème est encore plus prégnant dans le cas du multi-traitement, comme le contexte des moteurs d'offres, c'est-à-dire lorsque plusieurs actions sont possibles pour un individu. À notre connaissance, il n'existe pas de travaux sur l'impact des biais pour l'uplift dans le cas du multi-traitement. Pour résoudre cet écueil, nous proposons un protocole d'évaluation de l'uplift multi-traitement dans le cas du biais de non-random assignment. Muni de ce protocole, nous évaluons les performances des principales méthodes d'uplift multi-traitement de la littérature. Nous constatons que ces méthodes ont des comportements différents face au biais et nous en tirons des leçons.

1 Introduction

La modélisation de l'uplift est une approche prédictive qui vise à estimer l'impact de plusieurs traitements sur le comportement d'un individu. Ce type de modélisation est particulièrement utile dans divers domaines tels que la médecine personnalisée et la publicité. Dans le cadre des moteurs de recommandation d'offres marketing, la mesure de l'uplift est cruciale pour déterminer la campagne qui générera l'impact le plus favorable chez un client. Il ne s'agit pas seulement de prédire si une personne va répondre à une campagne, mais aussi d'estimer l'effet d'une campagne par rapport à d'autres. Ainsi, les modèles d'uplift aident à allouer de manière optimale les ressources marketing en ciblant les individus pour lesquels l'effet d'une campagne sera le plus significatif.

La littérature s'est intéressée initialement au cas du mono-traitement : seul l'impact d'un traitement est mesuré, par rapport au traitement de contrôle (ou non-traitement). On peut citer les arbres de décision (Radcliffe et Surry, 2011) ou les *Random Forest* (Guelman et al., 2015).

Cet article porte sur le cas le plus difficile du multi-traitement : l'objectif est d'ordonner et de quantifier l'impact de plusieurs traitements par rapport à un traitement de contrôle.

Les méthodes d'uplift reposent sur l'hypothèse que les différents groupes (traitements, contrôle) sont homogènes, c'est-à-dire qu'il n'y a pas de biais entre les individus des différents groupes. En pratique, les données sont de nature observationnelle. Dans un système de recommandation d'offres, l'attribution d'un traitement (par exemple une offre marketing) peut dépendre des caractéristiques de l'individu. En d'autres termes, certains clients ont une probabilité plus élevée d'être ciblés par certaines campagnes en raison de leur comportement passé ou d'autres facteurs, ce qui peut fausser l'estimation de l'uplift et limiter l'efficacité des moteurs de recommandation.

Il existe de multiples biais. Par exemple, la non réponse à un appel commercial introduit un biais entre personnes contactées et non contactées. Le but de cet article est d'étudier les comportements des principales méthodes d'uplift dans le cas du biais d'assignation non aléatoire *"Non-Random Assignment"* (Berk et al., 1988). Ce biais, noté *NRA*, est très courant dans les systèmes de recommandation. Il survient lorsque la probabilité de recevoir un traitement dépend des caractéristiques de l'individu. De façon générale, il existe peu de travaux sur l'étude des biais dans le contexte de l'uplift (Rafla et al., 2022) et, à notre connaissance, notre travail est le premier qui porte sur les méthodes multi-traitement.

Notre objectif est d'étudier comment les méthodes d'uplift se comportent face au biais *NRA* et à la réalité des problèmes rencontrés par les systèmes de recommandation. Ce contexte se caractérise par des données déséquilibrées, un grand nombre de traitements, et des situations où l'uplift est faiblement marqué. Pour répondre à cet objectif, nous traitons trois questions de recherche : comment les performances des méthodes d'uplift évoluent en fonction de l'importance du biais *NRA*, du nombre de traitements et du fait que l'uplift soit clairement marqué ou non (dans la suite, nous appelons *contraste* cette caractéristique). Pour répondre à ces questions, nous concevons un protocole expérimental apte à quantifier l'impact du biais *NRA* dans le cas du multi-traitement. Nous explicitons les propriétés que doit vérifier un tel protocole. Pour une évaluation exacte de l'uplift, nous montrons qu'il est nécessaire de passer par des données synthétiques et nous décrivons comment générer de telles données. Les résultats expérimentaux montrent que les méthodes ont des comportements différents en présence du biais. Nous tirons un certain nombre de leçons utiles de ces résultats pour concevoir dans le futur des techniques adaptées au cas du biais *NRA* dans le multi-traitement.

Cet article est organisé comme suit. La section 2 donne les définitions élémentaires et indique comment évaluer l'uplift. Nous décrivons notre protocole et la génération des données qui s'ensuit à la section 3. La section 4 présente les résultats expérimentaux permettant de mieux comprendre l'impact du biais *NRA*.

2 Contexte et préliminaire

2.1 Définition et notation

Uplift. L'uplift (Rubin, 1974; Radcliffe et Surry, 2011) est une estimation de l'impact causal d'un traitement spécifique sur un individu donné. Soit Y une variable exprimant l'impact d'un traitement (par exemple, dans le cas du marketing, "clique" et "ne clique pas") et T une variable de traitement. On note $Y_i(T = 1)$ le résultat de l'individu i lorsqu'il reçoit le traitement $T = 1$

et $Y_i(T = 0)$ pour le traitement $T = 0$. Formellement, l'uplift est une estimation de $Y_i(T = 1) - Y_i(T = 0)$. Si ce résultat est positif (resp. négatif), alors le traitement $T = 1$ a un impact positif (resp. négatif) pour cet individu.

La majorité de la littérature s'intéresse à l'uplift monotraitement où la variable de traitement T est binaire. Cette situation correspond à un groupe de *contrôle* et un groupe de *traitement*. $Y(T = 1)$ (resp. $Y(T = 0)$) est alors le résultat d'un individu lorsqu'il reçoit le traitement (resp. lorsqu'il ne le reçoit pas). Dans ce cadre, l'uplift est : $\tau(\mathbf{x_i}) = \mathbb{E}[Y|T = 1, X = \mathbf{x_i}] - \mathbb{E}[Y|T = 0, X = \mathbf{x_i}]$. Cependant, cette modélisation n'est plus valable s'il y a plusieurs traitements ($T \in \{0, 1, ..., K\}$). En effet, il n'y a plus un seul uplift (entre le traitement et le contrôle), mais K uplifts, à savoir un uplift entre chacun des traitements et le contrôle. L'uplift du traitement k est alors $\tau_k(\mathbf{x_i}) = \mathbb{E}[Y|T = k, X = \mathbf{x_i}] - \mathbb{E}[Y|T = 0, X = \mathbf{x_i}]$.

Politique de traitement. L'objectif de l'estimation de l'uplift multi-traitement est de déterminer le traitement avec le meilleur impact, c'est-à-dire le traitement ayant l'uplift le plus élevé. Le traitement idéal pour l'individu $\mathbf{x_i}$ est donc : $\pi(\mathbf{x_i}) = \arg\max_{k=0,1,...,K} \tau_k(\mathbf{x_i})$. Pour un système de recommandation, l'uplift est une mesure évaluant l'effet de chaque traitement sur les individus, permettant de déterminer la meilleure offre pour chaque client.

2.2 Modélisation de l'uplift multi-traitement

Nous présentons les principales méthodes d'uplift selon deux grandes familles, les *meta learners* et les méthodes relevant de l'approche directe.

Meta Learners. Les *meta learners* sont très classiques (Vanschoren, 2019). Ces approches utilisent des modèles usuels d'apprentissage automatique pour estimer l'uplift. Les *S Learners* apprennent un modèle unique en incluant le traitement comme une variable supplémentaire dans les prédictions. Les *T Learners* construisent un modèle distinct pour chaque groupe de traitement, estimant ainsi l'uplift en comparant les résultats entre les groupes. Le *X Learner* est une extension du *T Learner*. La différence est qu'il applique les estimateurs sur un changement de variable, qu'il pondère à l'aide du score de propensité (Rubin, 2001). Le *R Learner* estime l'uplift à l'aide d'une fonction de *loss*, et en excluant certains individus. Enfin, le *DR Learner* est une approche combinant le *T Learner* et le score de propensité.

Approches directes. Le principe des méthodes de cette famille consiste à modéliser directement l'uplift en adaptant des méthodes d'apprentissage supervisé classiques pour les rendre appropriées à la problématique de la modélisation de l'uplift. Dans cet article, nous utilisons des méthodes fondées sur les arbres de décision : *ED*, *Chi* et *CTS* (Rzepakowski et Jaroszewicz, 2012; Zhao et al., 2017).

2.3 Evaluation de l'uplift

Une difficulté intrinsèque à la modélisation de l'uplift est qu'il est impossible pour un individu de savoir si le traitement choisi est optimal car ses réponses aux traitements alternatifs ne peuvent pas être observées. Aussi, les méthodes classiques d'évaluation de l'apprentissage automatique ne peuvent pas être utilisées car cela impliquerait de comparer des résultats contrefactuels, c'est-à-dire de comprendre ce qui serait arrivé si un individu traité n'avait pas reçu de traitement, ou inversement. C'est ce que l'on appelle le problème fondamental de l'inférence causale (Rubin, 1974). Cependant, l'uplift d'un individu peut être estimé empiriquement en

considérant et comparant deux groupes d'individus : un ensemble d'individus ayant reçu le traitement et un ensemble d'individus n'ayant pas reçu le traitement.

Expected Outcome. Bien qu'il ne soit pas possible avec des données observationnelles de connaître l'uplift exact, il est cependant possible de comparer des politiques entre elles grâce à l'*expected outcome* (Zhao et al., 2017). L'*expected outcome* est une mesure utilisée pour évaluer l'effet moyen d'une politique de traitement π donnée sur une population. Le principe est d'estimer quel aurait été le gain (i.e. la valeur de Y) si telle politique avait été suivie. À l'aide d'une nouvelle variable aléatoire Z, respectant $\mathbb{E}[Z] = \mathbb{E}[Y|T = \pi(X)]$, l'*expected outcome* estime $E[Z]$ grâce aux données test, et donc $\mathbb{E}[Y|T = \pi(X)]$ (Zhao et al., 2017).

Root Mean Square Error (*RMSE*). Comme la section suivante va le montrer, nous allons utiliser dans cet article des données synthétiques afin de disposer des valeurs exactes de l'uplift pour chaque individu. Il est alors possible d'utiliser la *RMSE*. La *RMSE* est une mesure classique qui quantifie l'écart entre les valeurs prédites et les valeurs observées. Pour un modèle de prédiction d'uplift multi-traitement, la *RMSE* pour un individu $\mathbf{x_i}$ est définie par $\sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\hat{\tau}_k(\mathbf{x_i}) - \tau_k(\mathbf{x_i})\right)^2}$ où K est le nombre total de traitements, $\hat{\tau}_k(\mathbf{x_i})$ est l'estimation de l'uplift k pour l'individu $\mathbf{x_i}$ et $\tau_k(\mathbf{x_i})$ la valeur réelle de l'uplift k pour cet individu.

3 Protocole pour évaluer l'impact du biais *NRA* dans l'évaluation de l'uplift multi-traitement

3.1 Problématique

Le biais *NRA* est un enjeu majeur dans l'estimation de l'uplift (Berk et al., 1988), en particulier lorsqu'il s'agit de personnaliser des traitements dans des contextes complexes tels que les systèmes de recommandation d'offres. Le biais *NRA* survient lorsque la probabilité de recevoir un traitement dépend des caractéristiques de l'individu. Formellement, ce biais est défini lorsque $P(X|T = i) \neq P(X|T = j)$ avec $i \neq j$. Cette problématique est peu étudiée dans la littérature (Rafla et al., 2022) et pas encore dans le cas du multi-traitement.

Quelques travaux ont été réalisés pour comparer expérimentalement les techniques d'uplift existantes. Olaya et al. (2020) ont réalisé un benchmark de méthodes d'uplift dans le cadre du multi-traitement et ils ont montré l'intérêt du score de propensité (Rubin, 2001). Toujours dans le cas du multi-traitement, Gubela et al. (2024) comparent des méthodes d'uplift dont l'impact du traitement est continu (par exemple, un revenu). Cependant, aucun de ces travaux ne porte sur des données biaisées.

3.2 Conception du protocole

Le protocole doit évaluer l'impact du biais *NRA* et assurer une comparaison objective des différentes méthodes. Nous précisons maintenant comment les propriétés ci-dessous permettent de réaliser cette tâche.

Propriétés pour évaluer l'impact du biais. Pour évaluer l'impact du biais *NRA*, le protocole doit permettre l'ajout de biais dans les données tout en quantifiant le taux de biais, afin de contrôler son introduction et d'étudier comment les méthodes réagissent à des taux crois-

sants de biais. La complexité des données étant un paramètre d'étude, celle-ci est introduite en faisant varier le nombre de groupes de traitement et les valeurs d'uplift.

Propriétés pour assurer une comparaison objective des méthodes. Les méthodes doivent être comparées sur une base équitable. Pour cela, les données d'entraînement et de test doivent être identiques pour chacune des méthodes. Il faut aussi maintenir un même nombre d'individus pour chaque taux de biais introduit, afin que les différences de performance des modèles puissent être attribuées uniquement à la présence de biais, et non à un manque de données.

L'importance des données synthétiques. Pour évaluer les méthodes d'uplift, nous avons besoin de contrôler de façon précise les caractéristiques des populations, les traitements appliqués, les taux de biais, de moduler la complexité des données et de connaître la valeur exacte de l'uplift pour les individus. Générer des données synthétiques permet de répondre à ces attentes. La section 3.4 détaille comment nous générons les données.

3.3 Questions de recherche

Notre étude expérimentale [1] a pour but de répondre aux questions de recherche suivantes :

QR 1 : comment les performances des méthodes d'uplift évoluent-elles en fonction de l'importance du biais *NRA* ? Les performances des méthodes d'uplift se dégradent-elles à mesure que le taux de biais *NRA* augmente ? Si oui, cette dégradation a-t-elle un comportement linéaire ou bien observe-t-on un effondrement des performances à partir d'un certain taux de biais ?

QR 2 : quel est l'impact du nombre de traitements ? Par exemple, les performances des méthodes se dégradent-elles à mesure que le nombre de traitements augmente ?

QR 3 : quel est l'impact du contraste ? La différence d'uplift entre groupes peut être plus ou moins forte. Plus elle est forte, plus le problème est estimé facile et plus elle est faible plus le problème est jugé difficile. Ce paramètre est appelé *contraste* (cf. section 3.4). Le but est d'étudier les performances des méthodes en fonction du contraste de l'uplift.

3.4 Génération des données

Afin de créer une structure adaptée à notre étude, nous générons des données en deux dimensions suivant deux variables X_1 et X_2 :

- chaque individu $\mathbf{x_i}$ est décrit suivant ces deux variables. X_1 et X_2 sont générées indépendamment suivant une loi uniforme sur l'intervalle $[0, 1]$.
- chaque variable X_1 et X_2 est divisée en 10 intervalles égaux. Les deux variables forment ainsi une grille de 100 cellules. Chaque cellule représente un sous-groupe d'individus partageant des valeurs similaires pour X_1 et X_2.
- pour chaque cellule, une valeur uniforme de $\mathbb{E}[Y|X = \mathbf{x_i}, T = t_k]$ est attribuée à chacun des K traitements. Ces valeurs reflètent l'espérance du résultat pour les individus dans cette cellule lorsqu'ils reçoivent le traitement t_k.
- pour chaque individu, un traitement T est tiré aléatoirement parmi les K traitements disponibles. Une fois le traitement assigné, la variable Y est générée suivant une loi de Bernoulli $Y \sim B(\mathbb{E}[Y|X = \mathbf{x_i}, T = t_k])$, où $\mathbb{E}[Y|X = \mathbf{x_i}, T = t_k]$ est l'espérance conditionnelle du résultat pour l'individu donné et le traitement reçu.

1. https://github.com/Nathpreums/Uplift_Multitreatment_Benchmark_NRA

X_1	X_2	$\mathbb{E}[Y\|T=t_1,X]$	...	$\tau_{t_1,t_0}(X)$	...	T	Y
$\mathbf{x_{i,1}}$	$\mathbf{x_{i,2}}$	$\mathbb{E}[Y\|T=t_1,X=\mathbf{x_i}]$	...	$\tau_{t_1,t_0}(\mathbf{x_i})$	...	t	0 ou 1

TAB. 1 – *Tableau des valeurs générées de* $\mathbb{E}[Y|T,X]$ *et des* τ_{T,t_0} *pour un individu* $\mathbf{x_i}$

Le tableau 1 illustre le format des données. Nous modélisons le contraste de la façon suivante. Pour chaque cellule, nous définissons deux valeurs possibles pour l'espérance conditionnelle $\mathbb{E}[Y|T=t_k,X=\mathbf{x_i}]$ notées e_{Y_+} et e_{Y_-}. Ces valeurs représentent les espérances dans les cellules dites "positives" et "négatives", avec $e_{Y_+} > e_{Y_-}$. Nous avons fixé trois couples de valeurs pour e_{Y_+} et e_{Y_-} : $\{0,1\}, \{0.2,0.8\}, \{0.4,0.6\}$. Ces couples représentent les différents niveaux de contraste entre les espérances des cellules, du plus élevé ($\{0,1\}$) au plus faible ($\{0.4,0.6\}$).

Nous avons choisi les valeurs 2, 4, 8, 16 et 32 pour le nombre de traitements possibles, auxquels s'ajoute un groupe de contrôle (traitement $T=0$).

Nous avons fixé la proportion de cellules "positives" (cellules pour lesquelles $\mathbb{E}[Y|T=t_k,X=\mathbf{x_i}]=e_{Y_+}$) à 20% de l'ensemble des cellules. Cela permet de simuler des scénarios où seules certaines zones de l'espace des variables bénéficient réellement des traitements. Nous avons choisi d'utiliser 5000 individus par traitement dans les jeux de données d'entraînement. Pour générer le biais *NRA*, il est nécessaire de disposer d'au moins du double du nombre d'individus par groupe de traitement. Nous avons donc généré 10000 individus par traitement afin de pouvoir introduire des déséquilibres entre les groupes de traitement. Pour garantir la robustesse des résultats et obtenir des estimations fiables des performances des modèles, chaque jeu de données est généré 10 fois, soit un total de 150 jeux de données différents (3 couples de valeurs $\{e_{Y_+}, e_{Y_-}\}$, 5 valeurs pour le nombre de traitements, et 10 répétitions pour chaque configuration). En parallèle, nous avons généré des jeux de données de test indépendants des jeux d'entraînement, avec 15000 individus par groupe de traitement. Ces jeux de données permettent d'évaluer les modèles sur des échantillons non utilisés lors de l'entraînement.

3.5 Génération du biais *NRA*

Le biais *NRA* provient d'un déséquilibre entre les distributions des individus parmi les différents traitements. Pour induire ce biais dans nos données, nous procédons comme suit.

Création du motif de biais. Pour chaque traitement, 50% des cellules de la grille issue de la génération des données (soit 50 cellules) sont sélectionnées pour former le motif de biais.

Déséquilibre dans la distribution des individus. Afin d'introduire un déséquilibre entre les cellules biaisées et non biaisées, pour un taux de biais donné, nous modifions le nombre d'individus assignés aux cellules du motif de biais. Concrètement, pour un biais de $X\%$, le nombre d'individus présents dans les cellules du motif de biais est réduit de $X\%$, tout en augmentant proportionnellement le nombre d'individus dans les cellules en dehors de ce motif.

Interprétation des taux de biais. Un biais de 50% signifie qu'il y a 50% d'individus en moins dans les cellules appartenant au motif de biais par rapport à une situation sans biais (0% de biais). Ces individus sont redistribués dans les cellules qui ne font pas partie du motif de biais. À l'opposé, un biais de 0% correspond à une distribution équilibrée, où le nombre d'individus est identique entre les zones biaisées et non biaisées.

Ce mécanisme de génération permet de simuler de manière contrôlée différents taux de biais *NRA*.

4 Résultats expérimentaux

4.1 Expérimentation et résultats

4.1.1 Influence du biais (QR 1)

Les figures 1 et 2 illustrent respectivement la *RMSE* et l'*expected outcome* de l'ensemble des modèles en fonction du biais *NRA*. On constate clairement que ce biais affecte les performances globales de tous les modèles.

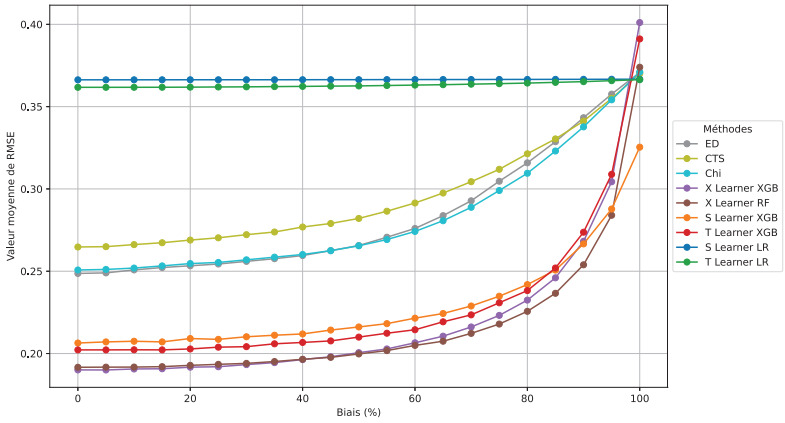

FIG. 1 – *Comparaison de la RMSE des modèles en fonction du biais NRA. Un point est obtenu en faisant la moyenne de la RMSE sur tous les jeux de données pour chaque nombre de traitements et niveau de contraste, à taux de biais fixé.*

Toutefois, en analysant les résultats modèle par modèle, des comportements distincts émergent. En particulier, les *meta learners* basés sur *XGBoost* (XGB) affichent de très bonnes performances, ainsi que le *X Learner* avec *Random Forest* (RF). En revanche, les modèles basés sur des forêts, comme les algorithmes *ED*, *Chi* et *CTS*, montrent des performances inférieures. Enfin, les autres *meta learners* utilisant la régression logistique sont significativement moins performants, car le modèle LR est probablement trop simpliste pour capturer la complexité des données. Ensuite, en ce qui concerne les modèles *R Learner* et *DR Learner*, bien que leurs performances soient médiocres selon la *RMSE*, ils se situent au même niveau que les forêts aléatoires lorsque l'on considère l'*expected outcome*. En effet, ces deux critères ne mettent pas en valeur les mêmes propriétés. La *RMSE* mesure la précision de l'estimation de l'uplift. Quant à l'*expected outcome*, il met en évidence le fait qu'un modèle arrive à bien ordonner ou non les traitements. On en déduit ainsi que les modèles *R Learner* et *DR Learner* prédisent mal les valeurs d'uplift, mais ordonnent mieux leurs prédictions. Concernant la sensibilité au biais, toutes les méthodes voient leurs performances chuter avec l'augmentation du biais *NRA*, à l'exception notable du *T Learner* et du *S Learner* basés sur la régression logistique. Ces modèles, trop simples pour capturer la complexité des données, sont moins affectés par le biais, car leurs performances se rapprochent déjà de celles d'une prédiction aléatoire. Parmi les autres

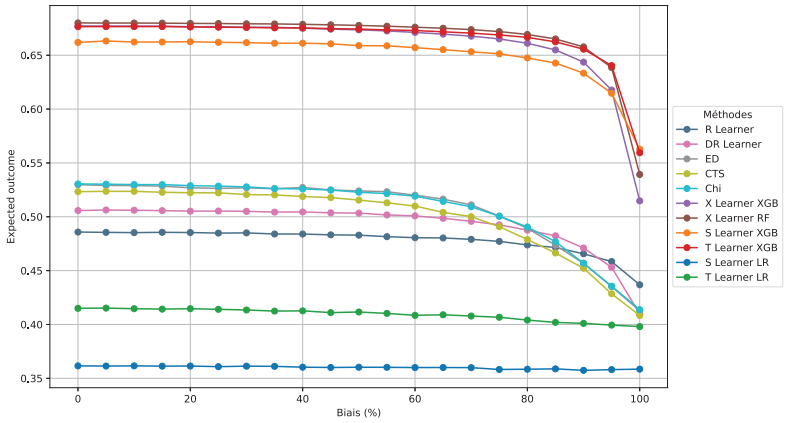

FIG. 2 – *Comparaison de l'expected outcome des modèles en fonction du biais NRA. Un point est obtenu en faisant la moyenne de l'expected outcome sur tous les jeux de données pour chaque nombre de traitements et niveau de contraste, à taux de biais fixé.*

modèles, la majorité semble avoir une résistance comparable au biais, bien que le *R Learner* montre une sensibilité légèrement moindre, de même que le *S Learner XGB*.

4.1.2 Influence du nombre de traitements (QR 2)

La *RMSE* est plus appropriée pour étudier l'influence du nombre de traitements. En effet, la valeur de l'*expected outcome* augmente naturellement avec le nombre de traitements, proportionnellement au nombre de cellules dites « positives ». Cela s'explique par le fait que plus le nombre de traitements augmente, plus il est probable qu'il existe un traitement avec un impact positif pour chaque individu.

L'analyse des résultats (figure 3) montre l'existence de deux types de comportements parmi les méthodes. Un premier groupe de méthodes présente des performances relativement constantes quel que soit le nombre de traitements. Ce groupe inclut les *T Learners* et les *X Learners*. Un second groupe de méthodes voit ses performances décliner au fur et à mesure que le nombre de traitements augmente. Ce groupe inclut notamment les méthodes de forêts (*ED, Chi, CTS*), les *S Learners*, ainsi que les *R Learners* et *DR Learners*.

Cette dégradation des performances pour le deuxième groupe s'explique par la complexité croissante du problème : plus il y a de traitements, plus il y a de valeurs possibles pour la variable de traitement, ce qui complique l'apprentissage pour ces méthodes. En revanche, les *T Learners* et *X Learners* sont peu affectés par cette augmentation du nombre de traitements. En effet, dans ces modèles, la variable de traitement est gérée de manière indépendante : un modèle distinct est construit pour chaque traitement, ce qui rend ces méthodes agnostiques à la variation du nombre de traitements. Nous pouvons enrichir notre analyse en considérant l'*expected outcome*, normalisé par rapport à l'*expected outcome optimal* pour chaque nombre de traitements. Les résultats montrent une tendance similaire pour tous les modèles, à l'exception du *DR Learner*. L'augmentation de la *RMSE* selon le nombre de traitements pour ce

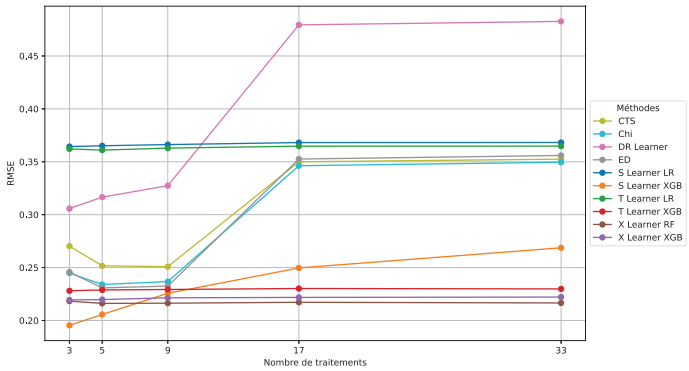

FIG. 3 – *Comparaison du RMSE des modèles en fonction du nombre de traitements. Un point est obtenu en faisant la moyenne de la RMSE sur tous les jeux de données avec toutes les valeurs de biais et tous les niveaux de contraste, à nombre de traitements fixé.*

modèle indique que les valeurs numériques d'uplift se détériorent progressivement. Cependant, l'*expected outcome* associé à ce modèle reste constant. Cela signifie que l'ordre des traitements suggérés par le modèle conserve des performances similaires, indépendamment du nombre de traitements considérés. Ainsi, pour le *DR Learner*, les recommandations de traitement ne semblent pas être sensibles à la variation du nombre de traitements.

Nous pouvons donc conclure que les *T Learners* et *X Learners* sont particulièrement bien adaptés aux problèmes où le nombre de traitements est élevé, contrairement aux autres méthodes qui voient leurs performances décliner dans ces situations. Le *DR Learner* ne se révèle pas efficace pour prédire l'uplift en présence d'un grand nombre de traitements, car les valeurs d'uplift estimées peuvent s'avérer peu fiables dans ce contexte. En revanche, dans des situations où seul le traitement recommandé est pertinent, tel qu'un système de recommandation d'offres, le *DR Learner* peut être considéré comme approprié.

4.1.3 Influence du contraste (QR 3)

Le contraste peut avoir un impact important sur les performances des modèles. Un faible contraste correspond à des uplifts moins marqués, ce qui rend plus difficile la distinction des traitements optimaux.

Cependant, l'analyse des résultats ne peut pas s'effectuer à partir des valeurs brutes de la *RMSE* et de l'*expected outcome*. En effet, la *RMSE* tend à diminuer lorsque le contraste diminue même à performances égales des modèles. Un phénomène similaire est observé pour l'*expected outcome*. Lorsque le contraste diminue, l'amplitude des résultats de l'*expected outcome* diminue également et les bonnes méthodes voient leur *expected outcome* baisser, tandis que les mauvaises méthodes ont un *expected outcome* qui augmente. Cela s'explique par le fait que, lorsque le contraste diminue, les bonnes décisions sont moins "récompensées" et les mauvaises décisions sont moins "pénalisées".

Aussi, afin d'analyser les résultats, nous effectuons une normalisation de l'*expected outcome*. Pour cela, nous considérons les valeurs de l'*expected outcome* du modèle optimal (mo-

dèle prédisant les traitements avec le meilleur impact) et du modèle aléatoire (modèle prédisant les traitements aléatoirement). Nous normalisons les résultats de sorte que l'*expected outcome* du modèle $Optimal$ soit égal à 1, et celui du modèle $Aléatoire$ soit égal à 0. La formule utilisée est : $EO_{Norm} = \frac{EO - EO_{Aléatoire}}{EO_{Optimal} - EO_{Aléatoire}}$. La figure 4 affiche les résultats avec les valeurs ainsi normalisées.

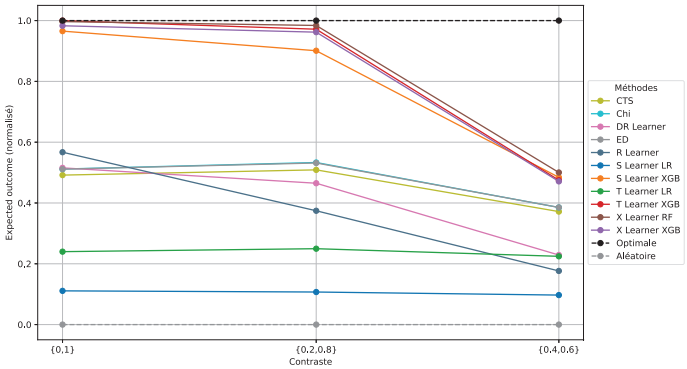

FIG. 4 – *Expected outcome normalisé pour les différentes méthodes. Un point est obtenu en faisant la moyenne de l'expected outcome sur tous les jeux de données pour chaque valeur de biais et nombre de traitements, à niveau de contraste fixé.*

L'analyse des résultats avec l'*expected outcome* montre une chute des performances liée au contraste pour les modèles *X Learner XGB*, *X Learner RF*, *T Learner XGB*, et *S Learner XGB*. Les forêts sont affectées, mais de manière moins prononcée. Toutefois, les méthodes qui étaient meilleures avec un contraste plus élevé le sont toujours avec un contraste plus faible.

4.2 Analyse et discussion

Résumé global des performances des modèles La figure 5 est un classement général des différentes méthodes sous la forme d'une étude de rang selon la $RMSE$ et l'*expected outcome*. Un point est obtenu en faisant la moyenne de la $RMSE$ ou de l'*expected outcome* sur tous les jeux de données, avec toutes les valeurs de biais et tous les niveaux de contraste. Les *meta-learners* basés sur *XGBoost* (*X*, *T* et *S Learner*), ainsi que le *X Learner RF*, se distinguent globalement comme les plus performants sur l'ensemble des jeux de données. Les modèles basés sur la régression logistique, bien que résistants au biais, présentent en réalité des performances bien plus faibles. Enfin, les méthodes de forêts aléatoires, comme *ED*, *Chi* et *CTS*, se montrent peu efficaces. Bien que ces méthodes soient robustes pour la prédiction d'uplift avec un faible nombre de traitements, leurs performances se détériorent rapidement lorsque le nombre de traitements augmente. Elles sont également sensibles aux variations de biais, ce qui en limite l'application dans des scénarios complexes, notamment de recommandation d'offres.

Leçons tirées des expérimentations Les résultats expérimentaux apportent plusieurs enseignements sur les modèles d'uplift multi-traitement en présence de biais. Les *meta-learners*

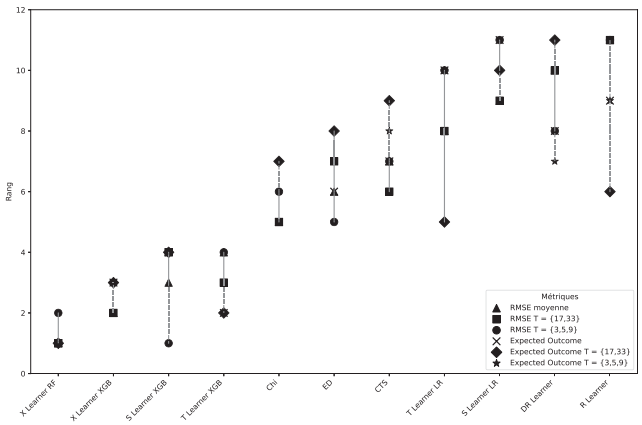

FIG. 5 – *Classement général des différents modèles en fonction des mesures de performance. Les formes indiquent la mesure de performance, tandis que les lignes montrent la dispersion du classement de chaque modèle en fonction des mesures de performance.* $RMSE\ T = \{3, 5, 9\}$ *(resp.* $RMSE\ T = \{17, 33\}$*) correspond à la* $RMSE$ *moyenne des méthodes sur les jeux de données avec 3, 5 ou 9 (resp. 17 ou 33) traitements. Expected Outcome* $T = \{3, 5, 9\}$ *(resp. Expected Outcome* $T = \{17, 33\}$*) correspond à l'expected outcome moyen des méthodes sur les jeux de données avec 3, 5 ou 9 (resp. 17 ou 33) traitements.*

basés sur *XGBoost* sont des candidats prometteurs pour des applications nécessitant la modélisation de nombreux traitements. Leur capacité à capturer des données non linéairement séparables tout en maintenant des performances robustes les rend particulièrement adaptés à des problèmes d'uplift complexes. Les modèles fondés sur la régression logistique (comme les *S* et *T Learners LR*) présentent une faible sensibilité aux biais et contraste, mais cela s'explique par leurs performances médiocres. Pour des applications où les données sont d'une plus grande complexité, les méthodes utilisant des modèles plus sophistiqués comme *XGBoost* devraient être privilégiées. Les méthodes de forêts aléatoires ne sont pas bien adaptées pour des scénarios avec un grand nombre de traitements. La tendance observée est que ces méthodes deviennent inefficaces dès que le nombre de traitements dépasse un certain seuil. Enfin, certains modèles ont de bonnes performances selon l'*expected outcome*, mais pas selon la *RMSE* ; l'ordre des traitements prédits est bon, mais pas l'estimation de l'uplift. La calibration est une piste de recherche pour améliorer cette estimation.

5 Conclusion et perspectives

Dans cet article, nous avons étudié l'impact du biais *NRA* sur les performances de modèles d'uplift multi-traitement de l'état de l'art. À notre connaissance, il s'agit de la première étude portant sur les effets d'un biais dans les modèles d'uplift multi-traitement. Nous avons conçu un protocole expérimental permettant de générer des données synthétiques d'uplift et d'y incorporer un biais *NRA* tout en le quantifiant. Les résultats expérimentaux montrent que les méthodes d'uplift ont des comportements différents en présence du biais *NRA* et permettent de

dégager des méthodes prometteuses pour des applications nécessitant la modélisation de nombreux traitements. Cette meilleure compréhension des comportements des méthodes d'uplift va permettre de définir des techniques pour résister au biais *NRA*. Il sera aussi intéressant d'évaluer ces modèles dans un contexte réel, comme les systèmes de recommandation d'offres.

Références

Berk, R. A., G. K. Smyth, et L. W. Sherman (1988). When random assignment fails : Some lessons from the minneapolis spouse abuse experiment. *Journal of Quantitative Criminology 4*(3), 209–223.

Gubela, R. M., S. Lessmann, et B. Stöcker (2024). Multiple treatment modeling for target marketing campaigns : A large-scale benchmark study. *Inf. Syst. Frontiers 26*(3), 875–898.

Guelman, L., M. Guillén, et A. M. Pérez-Marín (2015). Uplift random forests. *Cybernetics and Systems 46*(3-4), 230–248.

Olaya, D., K. Coussement, et W. Verbeke (2020). A survey and benchmarking study of multi-treatment uplift modeling. *Data Mining and Knowledge Discovery 34*, 273–308.

Radcliffe, N. J. et P. D. Surry (2011). Real-world uplift modelling with significance-based uplift trees. *White Paper TR-2011-1, Stochastic Solutions*, 1–33.

Rafla, M., N. Voisine, et B. Crémilleux (2022). Evaluation of uplift models with non-random assignment bias. In *20th Int. Symposium on Intelligent Data Analysis, IDA 2022, Rennes, France, April 20-22*, pp. 251–263. Springer.

Rubin, D. (2001). Using propensity scores to help design observational studies : Application to the tobacco litigation. *Health Services and Outcomes Research Methodology 2*, 169–188, doi: 10.1023/A:1020363010465.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology 66*(5), 688.

Rzepakowski, P. et S. Jaroszewicz (2012). Decision trees for uplift modeling with single and multiple treatments. *Knowl. Inf. Syst. 32*(2), 303–327.

Vanschoren, J. (2019). Meta-learning. *Automated machine learning : methods, systems, challenges*, 35–61.

Zhao, Y., X. Fang, et D. Simchi-Levi (2017). Uplift modeling with multiple treatments and general response types. *CoRR abs/1705.08492*.

Summary

This work addresses the challenge of uplift estimation in multi-treatment scenarios on biased data, such as recommendation systems. We propose an evaluation protocol to measure the impact of non-random assignment bias on multi-treatment uplift methods and analyze their performances. The results indicate different behaviors of uplift models leading to several messages to deal with biased data.