

Causalité et apprentissage automatique

Emmanuelle Claeys, Myriam Maumy-Bertrand,
Léon Bottou

13 mai 2020

1 Introduction : causalité et apprentissage automatique

Ce chapitre propose un retour d'expériences industrielles réalisées dans (Bottou et al., 2013) pour montrer comment la causalité affecte une application de *machine learning* à très grande échelle, où l'établissement d'un modèle causal peut se différencier d'une observation habituelle des données. Avec différents modèles, notamment celui de Robins et celui de Pearl (Robins, 1987; Pearl, 2009), on dispose maintenant de nouveaux outils, tels que les graphes causaux. Cependant, une attention particulière doit être portée sur les effets de la variance et les effets d'équilibres (nous définissons ici les effets d'équilibre avec l'analogie de l'équilibre physique où les variables causales se compensent). Nous présentons ici une méthodologie appliquée dans (Bottou et al., 2013) pour répondre aux problématiques suivantes :

- comment réduire les effets de la variance, afin de disposer d'estimateurs (comme le taux de clics) plus robustes (dans le sens statistique) ?
- Comment gérer les effets d'équilibre lorsque nous souhaitons vérifier des liens causaux entre les variables ?
- Quelles sont les approches permettant d'obtenir une intuition causale ?

Les deux premières parties s'appuieront sur une problématique de publicité *online*, et plus précisément pour des moteurs de recherche de type Google, Bing (qui en sont des cas particuliers (Varian, 2007; Edelman, Ostrovsky, & Schwarz, 2007)).

La section 2 présente le fonctionnement de la publicité *online* pour servir de cas illustratif. La section 3 présente le modèle causal et le rôle de l'échantillonnage. La section 4 montre comment intégrer un modèle causal lorsque nous sommes contraint à un système dynamique. La section 5 propose une série d'outils pour réduire la variance. La section 6 et la section 7 proposent des remarques sur la difficulté d'appliquer un tel modèle et la section 8 propose une approche permettant de déduire l'intuition causale. Enfin, la section 9 conclut sur les perspectives futures de la recherche en causalité.

2 La publicité *online* dans les moteurs de recherche

L'article récent de Nettles ([New Data Why People Hate Ads](#), s. d.)¹ fait le constat que les utilisateurs de moteurs de recherche n'aiment globalement pas la publicité ; trop présente, les utilisateurs se sentent poursuivis après avoir visité et/ou acheté un article, au sens général du terme, dans le passé. Cependant, le placement de la publicité *online* reste un investissement pour les sites d'e-commerce (*i.e.* : des annonceurs) de plus en plus conséquent, notamment grâce aux collectes de données massives. Nous pouvons nous demander pourquoi les annonceurs continuent d'adopter cette stratégie. Voici un exemple : un utilisateur cherche dans un moteur de recherche des lunettes. Il se peut qu'il soit déjà convaincu de la nécessité de son achat. L'annonceur souhaite que l'utilisateur voit des liens qui sont considérés comme de la publicité. L'utilisateur clique vers un lien publicitaire. Cependant, la publicité n'a pas convaincu l'utilisateur, elle lui a simplement fourni un lien vers lequel il aurait de toute façon cliqué. Pire encore, l'annonceur va continuer à envoyer ses publicités à l'utilisateur, puisqu'il a répondu à l'objectif de clic. Cet utilisateur sera soumis à des publicités de lunettes lors de ses prochaines connexions.

L'enquête de Nettles dit aussi que beaucoup de publicités sont jugées comme inappropriées par ceux qui les reçoivent. Dans ce cas il est naturel comprendre pourquoi l'annonceur les a proposées. Lorsque les annonceurs analysent leurs données, ces publicités reçoivent plus de clics que des publicités qui étaient réellement plus pertinentes pour les utilisateurs. Ce que les annonceurs omettent, ce sont les conséquences de ces clics. Au lieu d'un utilisateur heureux d'acheter le produit qu'il recherchait, ce dernier ne revient plus sur le site ou ignore toutes les publicités, qu'elles répondent ou non à leurs besoins. Ces exemples illustrent bien le problème de confusion causale auquel les annonceurs font face.

Un tel investissement montre que, quelque soit la robustesse d'un modèle de prédiction de clics, les utilisateurs ont pris l'habitude d'acheter sur internet. Un système d'enchères s'est mis en place ces dernières années, qui sera défini par la suite, fait intervenir trois acteurs :

- **L'utilisateur** : tape une requête et révèle l'objet de sa recherche par des mots-clés.
- **L'annonceur** : une compagnie marchande chargée de construire des publicités. Il doit sélectionner les bonnes opportunités, c'est à dire les requêtes en achetant des mot-clés présents dans la requête. Ces mots-clés sont vendus par le moteur de recherche qui les placera par la suite dans le résultat de la recherche formulée par l'utilisateur.
- **L'éditeur** : c'est le moteur de recherche. Il calcule des métriques suite à la recherche d'un utilisateur, fixe des enchères et choisit des publicités à lui montrer (ainsi que leur emplacement). Il doit également fixer un prix par clic (qui sera payé par l'annonceur si l'utilisateur clique éventuellement

1. <https://www.vieodesign.com/blog/new-data-why-people-hate-ads>

sur la publicité²).

Dans ce système, chaque acteur a des intérêts qui lui sont propres. Il est important de noter que le dernier objectif de l'éditeur est le plus compliqué. En effet, l'éditeur ne doit pas simplement maximiser son revenu mais également faire en sorte que l'utilisateur et l'annonceur soient satisfaits pour réinvestir à nouveau.

Le marché publicitaire total représente près d'un trillion de dollars par an, majoritairement diffusé par la télévision. Le marché de la publicité *online* aux états-unis représente 159 milliards de dollars et celles qui sont diffusées via les moteurs de recherche n'en représente qu'une petite partie. En revanche, les moteurs de recherche, qui doivent répondre à plus de 5,5 milliards de requêtes par jour (*Chiffres-clés de Google, s. d.*) ont l'avantage de savoir que l'utilisateur cherche quelque chose. La première étape pour l'annonceur est de déterminer si la requête d'un utilisateur est commerciale. La seconde est de proposer une publicité appropriée. Cependant, ces deux étapes font intervenir des problèmes causaux, sur lesquels nous nous concentrerons dans cet article.

2.1 Système des enchères publicitaires

Lorsqu'un utilisateur tape une requête dans un moteur de recherche, ce dernier est chargé de proposer des emplacements suite à la requête. Pour chaque emplacement disponible, des annonceurs enchérissent pour la place. La Figure 1 illustre une telle situation.

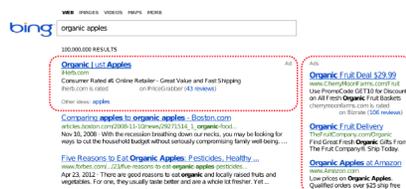


FIGURE 1 – Exemple dans un moteur de recherches d'annonces publicitaires

Dans ce système, interviennent des effets directs : si un éditeur fait augmenter le prix de l'enchère sans la rendre rentable, tous les annonceurs vont baisser leurs enchères et l'éditeur perdra de l'argent. Il y a également la présence d'effets indirects : quand un utilisateur clique, il achète éventuellement. Cette action a des conséquences sur ce qu'il verra dans le futur (par exemple lors de sa prochaine requête). Il peut également ne pas retourner sur ce moteur de recherche, ou ignorer les publicités proposées. L'annonceur, quant à lui, espère que le prix payé pour avoir placé sa publicité va aboutir à un achat. S'il n'obtient jamais d'achat, il va baisser ses enchères ou même ne plus mettre de publicités.

Le modèle de prédiction de clics va intégrer ces effets immédiats et secondaires, et ceci sur plusieurs semaines ou sur plusieurs mois.

2. Notons que le prix par clic n'est pas nécessairement égal au prix de l'enchère

Théorie des enchères : pour préciser les idées, nous allons présenter dans le paragraphe suivant le modèle global des enchères (Myerson, 1981; Milgrom, 2004). Un vendeur a un objet à vendre et plusieurs acheteurs peuvent enchérir. Chaque acheteur a une estimation différente de la valeur de cet objet et sait comment le système d'enchères fonctionne. Un acheteur annonce son offre en fonction de cette estimation, mais également en fonction de ce qu'il croit que l'objet peut valoir pour les autres. Le mécanisme de l'enchère prend en compte toutes les offres et détermine deux choses : qui est le vainqueur et ce qu'il doit payer.

Dans une enchère simple, le vainqueur sera celui qui propose la meilleure offre et il payera le prix annoncé. Ce système présente des failles : si un acheteur souhaite remporter l'enchère ; les acheteurs ont tout intérêt à baisser le plus possible la valeur estimée de cet objet pour maximiser le surplus. A *contrario* une enchère achetée à un prix trop élevé n'aura aucun intérêt (malédiction du vainqueur). Si les acheteurs font une proposition inférieure à la valeur réelle de l'objet, ils font un compromis entre «perdre l'enchère face à une offre plus élevée» et «payer moins que la valeur réelle».

Un deuxième système est plus intéressant : dans une enchère dite «de second prix» (Varian, 2007; Edelman et al., 2007), le vainqueur reste celui ayant fait la meilleure offre, mais il ne paie que l'offre de l'acheteur qu'il vient de dépasser, c'est à dire la second meilleure offre. Dans ce cas, on peut montrer qu'un acheteur a tout intérêt à faire une offre qui correspond exactement à son évaluation de l'objet (Riley & Samuelson, 1981) car il ne paiera que l'offre qui arrivera en second. Ce mécanisme d'enchère est optimal dans le sens qu'il maximise le surplus moyen de tous les participants. Le vendeur peut aussi établir un prix de réserve, c'est à dire faire lui même une offre. Seuls les acheteurs qui font une offre supérieure seront pris en compte. Comme le vendeur ne doit rien payer s'il gagne l'enchère, il peut donc établir un prix de réserve supérieur à son évaluation de l'objet. Cela biaise l'enchère en faveur du vendeur et au détriment du bénéfice de tous.

Application à la publicité en ligne : Soit $K \in \mathbb{N}$, nous définissons par q_1, q_2, \dots, q_K les K probabilités de clics de publicités a_1, a_2, \dots, a_K . La probabilité de clic q_i dépend d'une requête (composée de mot-clés) notée x (informations relatives à l'utilisateur, temps etc. . .), d'une publicité a et d'une position p_i (premier résultat affiché sous la barre de recherche, second, etc), choisie par l'annonceur. Le placement optimal consiste à trouver la position (p_i) pour une offre d'enchère b_i , où $i \in \{1 \dots K\}$. Nous cherchons à maximiser $\sum_i b_i q_i(x, a, p_i)$, c'est à dire la somme, sur toutes les positions, des offres d'enchères. Il est à noter que beaucoup de systèmes d'enchères fonctionnent sur ce principe, mais ils sont peu représentatifs de la réalité. Premièrement, les annonceurs donnent beaucoup de mots-clés pour lesquels ils sont seuls à enchérir (30% des cas). Deuxièmement, la théorie des enchères maximales fonctionne quand il n'y a qu'une seule enchère. Dans notre contexte, si le vendeur augmente sa part de bénéfice, les acheteurs

peuvent avoir recours à d'autres vendeurs. Troisièmement : les annonceurs n'enchérissent pas pour une requête spécifique. Ils réalisent des enchères multiples sur un temps défini (les acheteurs les plus actifs changent d'enchères toutes les quinze minutes). Enfin, les décisions de placement de publicité ont un impact sur l'utilisateur. Par exemple, les éditeurs qui envoient leurs utilisateurs vers des sites frauduleux ne sont pas viables à long terme puisque les utilisateurs ne reviendront plus. L'éditeur doit donc se poser la question suivante : quelles sont les publicités acceptables pour une requête x ? La probabilité de clics (q_i) et la notion de *rank score* (r_i) sont définis par :

$$q_i(x, a_i, p_i) = \underbrace{\beta(x, a_i)}_{\text{effet publicité}} * \underbrace{\gamma(x, p_i)}_{\text{effet position}} \quad (1)$$

$$r_i(x, a_i) = \underbrace{b_i}_{\text{offre}} * \underbrace{\beta(x, a_i)}_{\text{effet publicité}} \quad (2)$$

Après avoir calculé q_i et r_i via l'équation (1) et (2), l'éditeur assigne les publicités au rang r_i le plus élevé aux positions qui maximisent q_i . Le publicitaire paye ensuite la plus petite enchère qui lui aurait garanti la position la plus rentable. Bien qu'il semble prometteur en théorie, ce système ne fonctionne pas en pratique. En effet q_i ne représente plus la probabilité de clics car nous essayons de modéliser la satisfaction de l'utilisateur. Le prix de l'enchère n'est pas représentatif car les annonceurs ne paient pas leur valeur (ce n'est pas le prix optimal). Enfin, il y a généralement de nombreux rabais pour certaines enchères, qui aboutissent à un prix plus bas que le prix de réserve.

2.2 L'ingénierie des moteurs de placement

Pour répondre à un milliard de requêtes par jour, l'éditeur doit répondre à trois objectifs :

- sélectionner les publicités acceptables pour une requête x (section parmi quelques millions)
- sélectionner les q_i et les r_i via des modèles statistiques.
- intégrer les deux objectifs précédents dans le système pour sélectionner des publicités.

Pour répondre à ces trois objectifs, une analyse *offline* est réalisée par l'éditeur pour sélectionner un modèle statistique en utilisant des enregistrements d'activités d'utilisateurs (*i.e* : *des log*). Ce genre d'approche peut permettre de répondre aux questions telles que «Est ce que connaître la probabilité de clics sur la première publicité améliore le modèle de prédiction de probabilité de clics de la deuxième publicité?». Cette question sous-entend que la probabilité de clics de la première publicité affecte la deuxième. Une approche naïve est de regarder les données acquises sur des pages ayant eu au moins deux publicités et les partager en deux groupes. Ces deux groupes peuvent être des requêtes commerciales et des requêtes non commerciales; les requêtes commerciales ayant une probabilité de clics plus élevée. En effet, si par exemple un utilisateur tape

«Journée d'études statistiques» sur un moteur de recherches, il est peu probable qu'il clique sur une publicité. En revanche, s'il tape «Assurance automobile» le taux de clics de l'utilisateur est plus élevé, car il cherche quelque chose. Un exemple de résultats est rapporté dans la Table 2. Il semblerait que lorsque q_1 est une requête commerciale (probabilité de clics élevée), q_2 soit également une probabilité de clics élevée. Lorsque nous affichons une publicité, faut-il alors nécessairement en mettre une en dessous ? Pas forcément ! Une mauvaise observation des résultats peut masquer une variable de confusion causale. Il s'agit ici de la probabilité de clic que q_2 . La table 3 présente les observations sous ces contraintes. Ces résultats illustrent le paradoxe de Simpson (Simpson, 1951) : quelle que soit la valeur q_2 , cette valeur est toujours plus élevée quand la probabilité q_1 est faible (non commercial). Ce type de paradoxe peut apparaître lorsque les variables causales ne sont pas étudiées, et conduire à de mauvaises décisions. Dans ce cas d'exemple, si une publicité est très attractive, elle fera baisser le taux de clics de celle en dessous, et bien que cette hypothèse soit raisonnable, les données de la table 2 semblent dire le contraire.

Pour éviter cette situation, les entreprises issues du web procèdent généralement à de l'A/B test qui consiste à faire une expérience randomisée. Deux variantes du moteur de recherche sont programmées. Les utilisateurs sont affectés aléatoirement, soit à la variante A soit à la variante B. Après une collecte de données durant plusieurs mois, l'éditeur mesure différentes métriques pour décider quelle est la variante optimale (Blake, Nosko, & Tadelis, 2015).

Ces métriques peuvent être quantitatives : nombre moyen de publicités par page, nombre moyen de publicités au-dessus de la barre de recherche, nombre de clics par page, ...

Cependant, l'A/B test a ses limites : les tests randomisés supposent que l'issue de chaque expérience n'a pas d'impact sur les suivantes (hypothèses d'échantillons indépendants et aléatoirement sélectionnés) (Tang, Agarwal, O'Brien, & Meyer, 2010). Or, ici, le placement de publicités affecte à la fois l'utilisateur mais aussi l'annonceur. Ces placements vont impacter les métriques de performance dans les deux groupes d'utilisateurs. Pour contrer cet effet, il est possible de réaliser un A/B test sur un petit groupe d'utilisateurs (Claeys, Gancarski, Maumy-Bertrand, & Wassner, 2017), au prix d'une statistique moins fiable. Le modèle prédictif ne sera pas robuste.

Les métriques observées peuvent également complexifier le problème. Par exemple si l'éditeur veut optimiser le revenu par page, cela se fera au détriment de l'expérience utilisateur. S'il augmente les prix de réserve, il générera plus de revenus à court terme, mais perdra des annonceurs à long terme. Il doit donc chercher à considérer plusieurs métriques s'il veut améliorer son modèle. Il peut par exemple s'intéresser au taux de retour (est-ce qu'un utilisateur revient sur sa page ?), le nombre moyen de clics pondéré par le *lead scoring*³. Idéalement, l'édi-

3. Le lead scoring est un anglicisme qui désigne la pratique de calcul et d'affectation d'un score aux *leads* (contacts commerciaux ou prospects qualifiés) de l'entreprise.

teur voudrait entraîner un système sur plusieurs métriques. En pratique, c'est difficilement réalisable et donc peu utilisé. En effet, ces systèmes étant très complexes, il est difficile de leur donner un but précis. Par exemple, si une équipe fait de la prédiction de clics, nous pouvons penser que si les prédictions sont fiables, l'équipe chargée des enchères fonctionnera bien. La chaîne de développement est alors partagée en plusieurs entités distinctes et nous nous intéressons au résultat final. Mais il s'agit d'un modèle statistique et non d'un développement logiciel. Chaque équipe entraîne des erreurs d'estimation sur la probabilité de clic et ces erreurs auront un impact sur les composants du système d'enchères. Le résultat observé est inévitablement mauvais. La Table [1](#) résume les différences entre la théorie et la pratique

TABLE 1 – Placement optimal théorique et pratique

<p><u>Théorique :</u></p> <ul style="list-style-type: none"> — Sélection des publicités : (a_1, \dots, a_K) — Calcul des probabilités de clics : $q_i(x, a, p)$ — Placement qui maximise la valeur supposée de l'objet de l'enchère : $\max \sum_i b_i q_i(x, a_i, p_i)$. — Contrôle du prix des annonceurs par rapport au prix de réserve R : $b_i \cdot q_i(x, a_i, p_i) \geq R$. 	<p><u>Pratique :</u></p> <ul style="list-style-type: none"> — Séparation de l'effet publicité $\beta(x, a_i)$ et l'effet position $\gamma(x, p_i)$. — Sélection des publicités : (a_1, \dots, a_K). — Calcul des probabilités de clics : $q_i(x, a, p) = \beta(x, a_i) \cdot \gamma(x, p_i)$ — Choix de la publicité qui maximise $b_i \cdot \beta(x, a_i)$. — Choix de la position selon le <i>rank score</i> maximal $r_i(x, a_i) = b_i \beta(x, a_i)$ — Contrôle du prix des annonceurs par rapport au prix de réserve R : $\beta(x, a_i) \cdot \gamma(x, p_i) \geq R$.
--	--

TABLE 2 – Probabilité de clics de la deuxième publicité affichée (q_2) selon la probabilité de la première publicité (q_1)

	q_2
q_1 non commerciale	124/2000 6,2%
q_1 commerciale	149/2000 7,5%

TABLE 3 – Fréquences observées de clics de la deuxième publicité affichée (q_2) selon la probabilité de la première (q_1) illustrant le paradoxe de Simpson

	totale q_2	q_2 commerciale	q_2 non commerciale
q_1 non commerciale	124/2000 6,2%	92/1823 5,1%	32/177 18,1%
q_1 commerciale	149/2000 7,5%	71/1500 4,8%	78/500 15,6%

2.3 Modélisation de la log-probabilité de clic

Pour modéliser la log-probabilité de clics, nous pouvons utiliser la régression linéaire ordinaire avec une fonction de lien précisant une probabilité, où la variable expliquée est une action de clic ou non clic, et correspondant à une variable aléatoire suivant une loi de Bernoulli. Les $z_j \in \mathbb{R}$ sont les *features* associées au triplet (x, a_i, p_i) et servent de covariable dans le modèle linéaire $f(z)$ produisant une probabilité de clic q_i :

$$q_i(x, a_i, p_i) = s(f(z_i)) = \frac{1}{1 + e^{-f(z_i)}}. \quad (3)$$

En accord avec l'équation (3), q_i est également le produit de l'effet publicité $\beta(x, a_i)$ et de l'effet position $\gamma(x, p_i)$. L'approximation de q_i par $s(f(z_i))$ n'est pas exactement identique à $\beta(x, a_i) * \gamma(x, p_i)$ mais nous approximerons pour n observations de z :

$$q_i(x, a_i, p_i) = s\left(\sum_{j=1}^n w_j z_{j,i}\right) = s\left(\underbrace{w_n z_{n,i} + \dots + \dots}_{\text{effet publicité}} + \underbrace{\dots + w_1 z_{1,i}}_{\text{effet position}}\right). \quad (4)$$

Les poids w_j représentent l'effet publicité ou position selon les variables observées. Il est important de noter que q_i ne dépend pas de la valeur de l'offre d'enchère b_i dans ce modèle.

Limites : L'éditeur peut être amené à vouloir améliorer son modèle q présenté ci-dessus sans vouloir améliorer son système de placement (c'est à dire le système choisissant une position p). Il peut par exemple se demander quelles sont les variables qui améliorent son modèle q_i . Malheureusement, c'est généralement la position p_i qui fait varier la valeur q_i (comme dans l'exemple présenté en section 2.2 où la publicité affichée en premier, impacte la deuxième). Les données des moteurs de recherche ont montré qu'une variable de contexte x (exemple : jour/nuit pendant la requête), ou de pertinence du message publicitaire a (exemple : commercial non commercial), impactait peu la valeur de q_i . En revanche, la position p est fortement corrélée à q et faire varier la position p entraîne une forte variation sur la probabilité q . L'éditeur se retrouve face au dilemme suivant :

- S’il fait varier la variable position p , il trouvera les positions qui maximisent q mais ne saura pas comment sélectionner les publicités a par rapport à une requête.
- S’il enrichit son contexte (x) avec des variables supplémentaires il trouvera les profils d’utilisateurs qui maximisent q mais ne saura pas non plus sélectionner les messages a_i pour une requête.
- S’il veut observer q selon a , il saura quelle sont les publicités qui maximisent q mais ne saura prédire q pour une requête x .

Sa seule solution est de croiser différentes requêtes x avec différentes publicités a . Cependant, de telles observations entraînent une forte variance et rendent difficile l’estimation de la probabilité de clics q .

Conclusion : l’apprentissage mis bout à bout est difficile : l’éditeur ne dispose pas d’une méthode claire pour améliorer son système de position. La théorie des enchères suggère de former des modèles de probabilité de clic mais ne permet pas de répondre complètement à son problème. Des modèles de prédiction de clics ne permettent pas de savoir comment améliorer les paramètres du système causal (Sutton & Barto, 1988). Le paragraphe suivant illustre le concept de «trou noir» et montre comment la négligence de ces paramètres détériore ce système.

2.4 Risque du trou noir et stratégie d’exploration

Voici un exemple pédagogique illustrant le concept de trou noir. Nous nous intéressons à deux requêtes possibles

- x_1 : «diamant pas cher»
- x_2 : «google»

Nous considérons un nombre égal d’utilisateurs ayant tapé soit la requête x_1 soit la requête x_2 . Ces utilisateurs ont été soumis à l’une des trois requêtes possibles (en proportion égale)

- a_1 : une annonce pour une publicité de «diamant pas cher»
- a_2 : une annonce pour une publicité de «voiture pas chère»
- a_3 : une annonce pour une publicité de «bague de fiançailles»

Nous considérons dans cet exemple qu’une seule et unique publicité est montrée à un utilisateur et que toutes les offres d’enchères sont identiques. Les probabilités réelles de clic sont rapportées dans la Table 4

Modèle 1 : l’éditeur propose aux utilisateurs ayant tapé une requête x_1 ou x_2 , une annonce a choisi aléatoirement.

La probabilité de clics est égale à $\mathbb{P}[q] = \frac{1}{2}(\frac{7+2+9}{3} + \frac{2+2+2}{3}) = 0,04$

Modèle 2 : l’éditeur construit un modèle basé sur les mots en commun entre la requête x et l’annonce a . le système partage le trafic d’utilisateurs pour les répartir uniformément entre les annonces ayant au moins un mot-clé en commun. Si aucune annonce n’a de mot en commun, le trafic est réparti uniformément

TABLE 4 – Exemple jouet de fréquences de clics selon différents messages publicitaires

	a_1 «diamant pas cher»	a_2 «voiture pas chère»	a_3 «bague de fiançailles»
x_1 «diamant pas cher»	7%	2%	9%
x_2 «google»	2%	2%	2%

TABLE 5 – Probabilités réelles de clic selon différents messages publicitaires a après l'application du modèle 2

	a_1 «diamant pas cher»	a_2 «voiture pas cher»	a_3 «bague fiançailles»
x_1 «diamant pas cher»	7%	2%	?
x_2 «google»	2%	2%	2%

entre toutes les annonces. Les utilisateurs ayant tapé x_1 : «diamant pas cher» auront donc l'annonce a_1 : «diamant pas cher» ou l'annonce «voiture pas cher». Les utilisateurs ayant tapé x_2 : «google» verront l'une des trois annonces. Les probabilités de clics sont égales à

- $\mathbb{P}[q|x_1, a_1 \cup a_2] = \frac{7+2}{2} = 0,045$
- $\mathbb{P}[q|x_2, a_1 \cup a_2 \cup a_3] = \frac{2+2+2}{3} = 0,02$
- $\mathbb{P}[q] = \mathbb{P}[q|x_1, a_1 \cup a_2] + \mathbb{P}[q|x_2, a_1 \cup a_2 \cup a_3] = \frac{1}{2}(\frac{7+2}{2} + \frac{2+2+2}{3}) = 0,0325$.

La probabilité de clics totale est, avec un tel modèle, inférieure à celle obtenue par une affectation aléatoire des publicités aux utilisateurs. Pire encore ; ce modèle aura des conséquences sur les prochaines collectes de données. En effet, l'éditeur ne pourra qu'observer les données présentes dans la Table 5. Ce modèle simpliste qui n'utilise qu'une seule règle (mots-clés en commun) ne permet pas d'observer le cas (x_1, a_3) qui avait une probabilité de clic élevée (9%). Un trou noir est lorsqu'une probabilité conditionnelle devient non observable à cause du modèle choisi. Comme le montre l'exemple précédent, il peut entraîner une baisse de performance du modèle (ici une baisse de la valeur q). Les trous noirs peuvent apparaître dans de nombreuses situations, par exemple lorsqu'un annonceur se place en situation de monopole, ou lorsqu'une nouvelle publicité arrive sur le marché (probabilité de clics inconnue). Plus d'informations sont disponibles dans (Bottou et al., 2013). La seule solution pour éviter les trous noirs est d'utiliser une stratégie d'exploration : nous devons parfois afficher des publicités qui ne seraient affichées avec le modèle utilisé afin de pouvoir améliorer le modèle (2) plus tard.

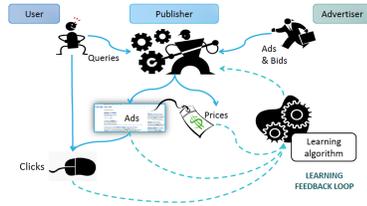


FIGURE 2 – Boucle de filtrage d'apprentissage

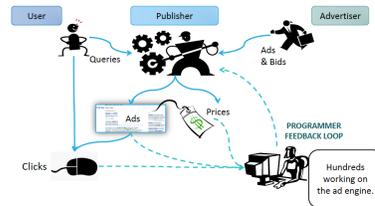


FIGURE 3 – Boucle de filtrage du programmeur

3 La boucle de filtrage

La Figure 2 représente la boucle de filtrage d'apprentissage et la Figure 3 représente une alternative possible : la boucle de filtrage du programmeur. Ils diffèrent de par l'intervention d'un agent automatique *versus* expert. Lorsque l'éditeur utilise une boucle de filtrage d'apprentissage, tout changement dans son algorithme de placement va affecter les données d'apprentissage, qui vont modifier le modèle. Sans surveillance sur le modèle, nous montrons de moins en moins de publicité selon certaines requêtes. La boucle de filtrage du programmeur est soumise au même risque, à la différence que ce sont des experts qui observent les données. Le problème causal est donc plus compliqué à étudier. Le schéma de la Figure 4 représente les acteurs du système.

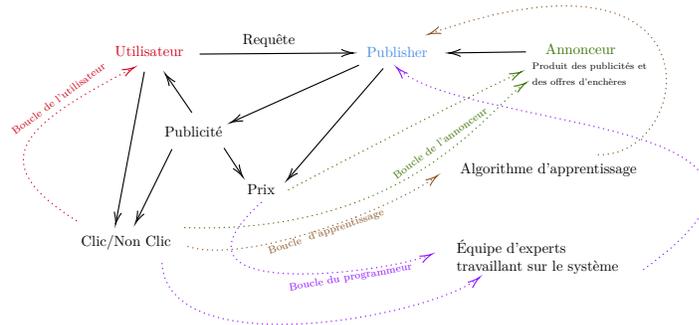


FIGURE 4 – Acteurs du système

Le graphe causal de la Figure 4 n'est pas acyclique : il contient en effet plus d'un cycle. Des travaux (Varian, 2007; Edelman et al., 2007) associés à la théorie des enchères résolvent bien la boucle de l'annonceur mais utilisent des hypothèses trop contraignantes :

- Il ne peut y avoir qu'une seule enchère (pas d'enchères multiples).

- Les probabilités de clics doivent être correctement estimées.
- Les publicités n’impactent pas le comportement futur des utilisateurs.
- Les publicitaires ne peuvent faire qu’une seule offre.

Il n’y a donc aucune façon claire de décrire l’ensemble du système comme des enchères issus d’une analyse théorique.

Des modèles de probabilité de clics utilisant la théorie des bandits à plusieurs bras (et extensions) (Auer, Cesa-Bianchi, & Fischer, 2002; Audibert, Munos, & Szepesvari, 2007; Li, Chu, Langford, & Schapire, 2010) se prêtent assez bien à la boucle de filtrage d’apprentissage. Ces modèles intègrent la notion d’exploration pour plus de performance, c’est à dire qu’ils proposent des configurations qui peuvent sembler sous-optimales. Cependant, ces modèles ignorent l’impact des publicités sur les utilisateurs, et l’impact des publicités sur les futures offres faites par les annonceurs. Il est difficile d’utiliser un ensemble de bandits pour modéliser l’impact de chaque acteur sur le système. Il faut donc commencer par représenter ce système par un graphique causal.

4 Modélisation du système causal

Nous définissons une partie du système via un modèle causal d’équations structurelles :S.E.M. (von Wright, 1971; Woodward, 2005). Ce modèle est représenté dans la figure 5. Les variables u et v sont inconnues, elles représentent la volonté de l’utilisateur et de l’annonceur. L’éditeur reçoit la requête x formulée par l’utilisateur via son moteur de recherche. Il génère ensuite une sélection s de publicités a possibles (en utilisant le score q , c’est-à-dire la prédiction de clic). Le score q lui permet également de fixer un prix c . Cette sélection s définit les pages vues par l’utilisateur. En fonction de cette sélection, l’utilisateur décide de cliquer ou de ne pas cliquer (y). L’action y va décider du revenu z que l’annonceur devra éventuellement payer. Chaque variable est définie par un système d’équations déterministes. Les variables connues sont cependant difficiles à connaître car elles changent dans le temps. Les ϵ caractérisent la part d’aléatoire dans l’expérience. Une **intervention** sur le système consiste à changer une équation. Par exemple, changer l’équation f_4 de scoring dans la Figure 5. Traditionnellement, si nous regardons les travaux de Pearl (Pearl, 2012), le changement d’équation consiste à fixer une constante. La variable u étant inconnue, il faut procéder à une **isolation** de x . Les variables exogènes suivent une distribution stationnaire : nous nions l’existence de la boucle de filtrage pour le moment. Sous de telles hypothèses, chaque variable du modèle causal de la figure 5 peut être calculée par des probabilités conditionnelles et représentée par un réseau bayésien (Pearl 1988). La figure 6 représente ce réseau.

L’intervention consiste à changer l’une des probabilités conditionnelles et observer son effet sur l’ensemble des variables impactées. Les «sous réseaux» bayésiens changent et nous observons ce qu’ils ont en commun. Le modèle de graphe causal permet de connaître, sous une intervention, quelles variables seront constantes et quelles variables vont changer. La question contrefactuelle est ici : «Quelles auraient-été les performances de mon système d’enchère, si

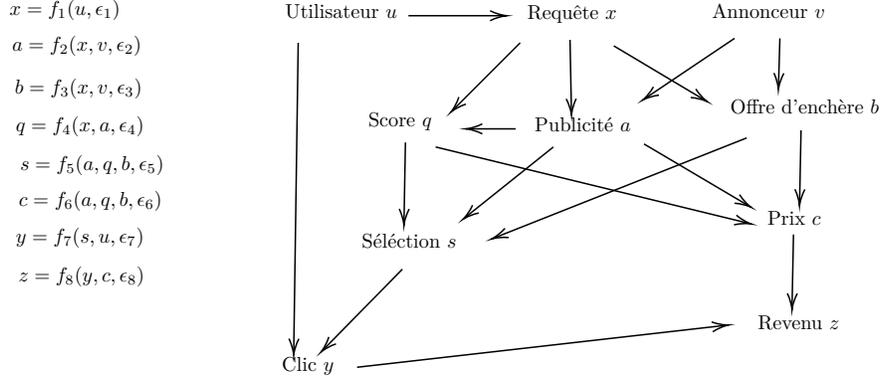


FIGURE 5 – Système causal

au moment de la collecte des données, nous avons utilisé un modèle causal q' plutôt que q ?».

Il faut collecter des données pour répondre à la question. Nous nous plaçons ici dans un cas différent de celui de Pearl (monde contrefactuel) de *potential outcomes*. Le *potential outcomes* consiste à observer les conséquences d'un changement q^* dans un modèle original (f_1, f_2, f_3, f_4) ou dans un modèle modifié $(f_1, f_2, f_3, f_4^*, f_5, f_6, f_7, f_8)$. Malheureusement, les métriques de performance du modèle de scoring ne permettent pas de rejouer l'expérience car les publicités et/ou les requêtes sont différentes.

4.1 Calcul de l'espérance contrefactuelle

Soit Y une variable d'intérêt, représentant l'espérance de clic y selon une configuration $\omega : Y = \mathbb{E}_{\omega \sim \mathbb{P}[\omega]}[y]$. ω représentant les valeurs u, v, x, a, b, q, s . Dans la suite du document, nous utiliserons uniquement la notation Y .

$$Y = \int_{\omega} y \mathbb{P}[\omega] d\omega.$$

Les modèles graphiques et d'équations structurelle proposés par Pearl, représenté dans la Figure 5. Pearl s'appuie sur une extension des modèles d'équations structurelles pour traiter les situations mettant en jeu des variables dichotomique et des dépendances non linéaires. Dans cette généralisation basée sur la simulation d'interventions hypothétiques dans le modèle, l'«effet» est défini non plus comme un coefficient dans une équation, mais comme une capacité à transmettre les changements entre les variable (Lecoutre, 2004). On définit une distribution altérée $\mathbb{P}^*[\omega]$ comme une factorisation markovienne telle que :

$$\begin{aligned} \mathbb{P}^*[\omega] &= \mathbb{P}[u, v] \mathbb{P}[x|u] \mathbb{P}[a|x, v] \mathbb{P}[b|x, v] \mathbb{P}^*[q|x, a] \\ &\times \mathbb{P}[s|a, q, b] \mathbb{P}[c|a, q, b] \mathbb{P}[y|s, u] \mathbb{P}[z|y, c]. \end{aligned} \quad (5)$$

$$P(\omega) = \mathbb{P}(u, v)$$

- * $\mathbb{P}[x|v]$
- * $\mathbb{P}[a|x, v]$
- * $\mathbb{P}[b|x, v]$
- * $\mathbb{P}[q|x, a]$
- * $\mathbb{P}[s|a, q, b]$
- * $\mathbb{P}[c|a, q, b]$
- * $\mathbb{P}[y|s, u]$
- * $\mathbb{P}[z|y, c]$

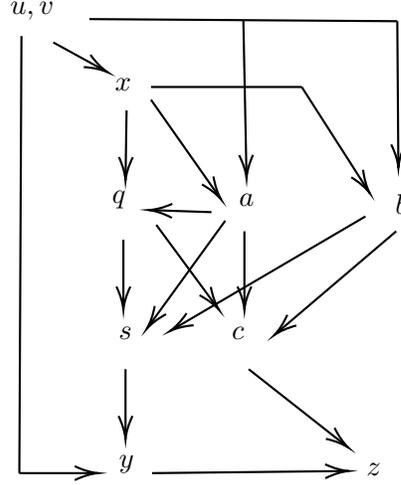


FIGURE 6 – Réseau bayésien

L'échantillonnage préférentiel est une méthode de réduction de la variance qui permet d'estimer l'espérance contrefactuelle. Reprenons Y une variable aléatoire d'intérêt i.i.d. (indépendante et identiquement distribuée), dépendante d'une quantité normalement distribuée $l(\omega)$ et d'une distribution altérée $\mathbb{P}^*[\omega]$. L'espérance de Y^* vaut :

$$\begin{aligned} Y^* &= \int_{\omega} l(\omega) \mathbb{P}^*[\omega] d\omega = \int_{\omega} l(\omega) \frac{\mathbb{P}^*[\omega]}{\mathbb{P}[\omega]} \mathbb{P}[\omega] d\omega \\ &\approx \frac{1}{n} \sum_{i=1}^n l(\omega) w_i \end{aligned} \quad (6)$$

où n est le nombre d'observations disponibles et w_i les poids définis par :

$$w_i = w(\omega_i) = \frac{\mathbb{P}^*[\omega_i]}{\mathbb{P}[\omega_i]} = \frac{\text{variables appartenant à } \mathbb{P}^*[\omega_i] \text{ mais pas à } \mathbb{P}[\omega_i]}{\text{variables appartenant à } \mathbb{P}[\omega_i] \text{ mais pas à } \mathbb{P}^*[\omega_i]}. \quad (7)$$

On considère par simplicité que $\forall x, a, \mathbb{P}[q|x, a] \neq 0$. Par exemple l'espérance de y sous une distribution altérée s'écrit :

$$\begin{aligned} Y^* &= \int_{\omega} y \mathbb{P}^*[\omega] d\omega = \int_{\omega} y \frac{\mathbb{P}^*[q|x, a]}{\mathbb{P}[q|x, a]} d\omega \\ &\approx \frac{1}{n} \sum_{i=1}^n y_i \frac{\mathbb{P}^*[q_i|x_i, a_i]}{\mathbb{P}[q_i|x_i, a_i]}. \end{aligned} \quad (8)$$

Le calcul des w_i peut devenir problématique lorsque les fonctions de $\mathbb{P}[\omega]$ et $\mathbb{P}^*[\omega]$ sont disjointes (voir la Figure 7). Nous préférons des distributions comme celles représentées sur la Figure 8. Un ratio élevé de $\frac{\mathbb{P}^*[q_i|x_i, a_i]}{\mathbb{P}[q_i|x_i, a_i]}$ nuisant à la

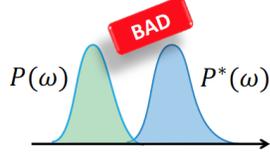


FIGURE 7 – Distribution disjointes de $\mathbb{P}[\omega]$ et de $\mathbb{P}^*[\omega]$

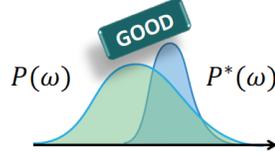


FIGURE 8 – Distribution conjointes de $\mathbb{P}[\omega]$ et de $\mathbb{P}^*[\omega]$

qualité de l'estimation de Y^* , il est possible d'utiliser des intervalles de confiance pour pallier ce problème. Comme pour une expérience aléatoire, nous intégrons une part d'aléatoire pour explorer suffisamment. Cependant, cette exploration sera active : son objectif sera de répondre à la question contrefactuelle. Nous souhaitons en effet avoir une probabilité non nulle d'effectuer l'action forcée f_4^* (voir Figure 5).

4.2 Calcul des intervalles de confiance

Le calcul de l'équation (6) via des outils classiques comme le théorème central limite est problématique car $w(\omega_i)$ prend des valeurs très grandes quand $\mathbb{P}[\omega_i]$ prend des valeurs très faibles (c'est à dire un petit nombre d'observations). Il faut donc définir une zone à explorer notée Ω_R , qui pour un ratio limite $R \in \mathbb{R}_*^+$ fixé :

$$\Omega_R = \{\omega : \mathbb{P}^*[\omega] < R\mathbb{P}[\omega]\}$$

le calcul de l'équation (6) dans Ω_R devient :

$$Y^* = \int_{\omega \in \Omega_R} l(\omega)\mathbb{P}^*[\omega]d\omega + \int_{\omega \in \Omega \setminus \Omega_R} l(\omega)\mathbb{P}^*[\omega]d\omega = \bar{Y}^* + (Y^* - \bar{Y}^*) \quad (9)$$

Nous définissons les poids $\bar{w}(\omega)$ (cf équation 7) permettant de considérer les $w(\omega)$ lorsque $\mathbb{P}^*[\omega] < R\mathbb{P}[\omega]$ tels que :

$$\bar{w}(\omega) = \begin{cases} w(\omega) & \text{si } \omega \in \Omega_R \\ 0 & \text{sinon} \end{cases} \quad (10)$$

L'estimation de \bar{Y}^* est plus facile à réaliser car elle intègre les poids $\bar{w}(\omega)$ bornés par R :

$$\begin{aligned} \bar{Y}^* &= \int_{\omega \in \Omega_R} l(\omega)\mathbb{P}^*[\omega]d\omega = \int_{\omega} l(\omega)\bar{w}(\omega)\mathbb{P}[\omega]d\omega \\ &\approx \hat{Y}^* = \frac{1}{n} \sum_{i=1}^n l(\omega_i)\bar{w}(\omega_i). \end{aligned} \quad (11)$$

Par son exploration limitée (Ω_R), l'estimateur \bar{Y}^* est biaisé. Avec $\omega \in \Omega \setminus \Omega_R$ représentant la zone non explorée, il est possible d'estimer ce biais par :

$$\begin{aligned}
0 \leq Y^* - \bar{Y}^* &\leq \int_{\omega \in \Omega \setminus \Omega_R} l(\omega) \mathbb{P}^*[\omega] d\omega \\
&\leq \mathbb{P}^*[\Omega \setminus \Omega_R] \max_{\omega \in \Omega \setminus \Omega_R} l(\omega) = (1 - \mathbb{P}[\Omega_R]) \max_{\omega \in \Omega \setminus \Omega_R} l(\omega) \\
&\leq (1 - \int_{\omega} \bar{w}(\omega) \mathbb{P}[\Omega_R] d\omega) \max_{\omega \in \Omega \setminus \Omega_R} l(\omega) \\
&\leq (1 - \frac{1}{n} \sum_{i=1}^n \bar{w}(\omega_i)) \max_{\omega \in \Omega \setminus \Omega_R} l(\omega). \tag{12}
\end{aligned}$$

Deux intervalles de confiance sont calculés par l'inégalité de Bernstein pour une valeur de R et un seuil de confiance δ choisis :

- l'équation (13) caractérise l'erreur ϵ_R autour de \bar{Y}^*
- l'équation (14) caractérise l'erreur ξ_R autour de Y^* .

$$\mathbb{P} \left[\hat{Y}^* - \epsilon_R \leq \bar{Y}^* \leq \hat{Y}^* + \epsilon_R \right] \geq 1 - \delta \tag{13}$$

$$\begin{aligned}
&\mathbb{P} \left[\bar{Y}^* - (1 - \frac{1}{n} \sum_{i=1}^n \bar{w}(\omega_i) + \xi_R) \max_{\omega \in \Omega \setminus \Omega_R} l(\omega) \right. \\
&\quad \left. \leq Y^* \leq \bar{Y}^* + (1 - \frac{1}{n} \sum_{i=1}^n \bar{w}(\omega_i) + \xi_R) \max_{\omega \in \Omega \setminus \Omega_R} l(\omega) \right] \geq 1 - \delta. \tag{14}
\end{aligned}$$

La différence $Y^* - \bar{Y}^*$ reflète le biais résultant d'un manque d'exploration. Par conséquent :

- L'intervalle de confiance de l'équation (13) représente l'incertitude qui résulte de la taille limitée de l'échantillon Ω_R . Un grand intervalle de confiance indique que l'échantillon est trop petit. Pour améliorer le résultat, il suffit de continuer à collecter des données en utilisant le même dispositif expérimental.
- L'intervalle de confiance de l'équation (14) témoigne de l'incertitude associée au domaine $\Omega \setminus \Omega_R$ insuffisamment exploré. Un intervalle de confiance important suggère d'améliorer l'estimation en utilisant des données particulières afin d'obtenir une meilleure couverture des fonctions contrefactuelles d'intérêt.

4.3 Expérimentation

Pour expérimenter l'approche présentée en section 4.2, nous nous intéressons au problème suivant : l'éditeur dispose de publicités à placer. Il doit décider du

nombre de publicités à positionner sur le résultat de la requête présentée à l'utilisateur (voir Figure 1). Nous nous intéresserons ici uniquement aux publicités placées au dessus des résultats du moteur de recherche (*Mainline*).

Nous définissons la *mainline reserve* comme le prix minimum fixé pour les publicités placées dans la *mainline*. La *mainline reserve* peut donc être vue comme un prix de réserve pour les publicités de la *mainline*. Ainsi, la *mainline reserve* détermine indirectement le nombre de publicité montrées au dessus des résultats de la recherche (*mainline reserve*). Ce nombre d'emplacements peut donc varier en fonction de l'intérêt commercial pour cette requête.

L'éditeur souhaite répondre à cette question contrefactuelle : «Quel aurait été le rendement du système de placement des publicités si nous avions augmenté le prix de la *mainline reserve* d'une valeur m (définie par la suite), sans susciter de réactions de la part des utilisateurs ou des annonceurs?».

Avant de déterminer la position p_i d'une publicité a_i , pour paramétrer la distribution conditionnelle des scores, une moyenne $\rho \in \mathbb{R}$ et une variance σ^2 sont choisies. Un nombre aléatoire ϵ est tiré selon une loi normale centrée réduite $\mathcal{N}(0; 1)$, et toute la *mainline reserve* est multipliée par m tel que $m = \rho e^{-\frac{\sigma^2}{2} + \sigma\epsilon}$. Ces coefficients m suivent une loi *log-normale* : $\ln \mathcal{N}(\mu; \sigma^2)$ avec $\mu = \frac{\sigma^2}{2} + \log(\rho)$. En rapportant la définition des poids w_i de l'équation (7), avec un des paramètres ρ^*, σ^* choisis :

$$w_i = w(\omega_i) = \frac{\mathbb{P}^*[w_i]}{\mathbb{P}[w_i]} = \frac{\int_{i=1}^n m_{i\rho^*,\sigma^*} dm_i}{\int_{i=1}^n m_{i\rho,\sigma} dm_i} = \frac{\mathbb{P}^*[q_i|m_{i\rho^*,\sigma^*}]}{\mathbb{P}[q_i|m_{i\rho,\sigma}]}. \quad (15)$$

Par exemple, nous tirons plusieurs coefficients en fixant les paramètres $\rho^* = 1$ et $\sigma^* = 0.3$. Avec ces paramètres, 95% de ces coefficients m_i sont dans l'intervalle $[0, 52; 1, 74]$.

Des observations décrivant 22 millions de pages de résultats de recherches ont été recueillies pendant cinq semaines consécutives. Nous faisons varier le paramètre ρ^* et gardons le paramètre σ^* à 0,3. La Figure 9 montre les résultats observés selon trois objectifs : (1) le nombre moyen de publicités par page affichée, (2) le nombre moyen de clics par page affichée (3) le revenu moyen par page affichée, en fonction de m^* . Les courbes montrent également des intervalles de confiance de l'équation (13) (non grisés) et de (14) (grisés) avec un intervalle à 95% autour des moyennes observées.

Ces courbes présentent les intervalles de confiance selon des paramètres que nous aurions observés si nous avions augmenté la *mainline* du pourcentage indiqué sur l'axe horizontal. Les zones pleine représentent les intervalles de confiance de l'équation (14) et les zones non remplies celles de l'équation (13).

Les carrés creux représentent les métriques mesurées sur les données expérimentales. Les cercles creux représentent des métriques mesurées sur une expérience aléatoire dont la *mainline* a été réduite de 18%. Les cercles remplis représentent les métriques effectivement mesurées sur un échantillon témoin fonctionnant sans aléatoire.

Les intervalles de confiance de l'équation (14) (grisés) augmentent fortement lorsque le paramètre ρ^* n'est plus observé dans les observations issues de la

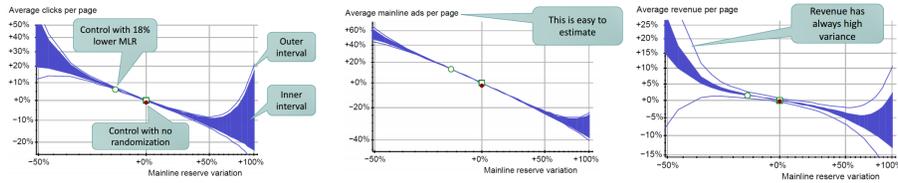


FIGURE 9 – Estimation des variations de trois paramètres de rendement en réponse aux changements ρ^* .

collecte de données. Le revenu moyen par page varie davantage car en pratique, seulement quelques requêtes très compétitives engendrent des prix d’enchère élevés et donc la majorité des revenus.

Afin de valider l’exactitude de ces estimations contrefactuelles, nous décidons de réduire de 18% la *mainline reserve* et affectons cette configuration à un nombre d’utilisateurs égal à celui utilisé pour l’expérience. Les cercles creux de la Figure 9 représentent donc les métriques réelles mesurées pendant la même période. Nous observons que les mesures réelles et les estimations contrefactuelles se correspondent.

Nous avons également voulu comparer ces résultats avec des publicités choisies aléatoirement et fait fonctionner le système de placement d’annonces non modifiées sur un nombre d’utilisateurs «témoin». Des publicités ”candidates” à une requête sont choisies aléatoirement et le nombre de publicités affichées ne change pas. Les cercles plein de la Figure 9 représentent les résultats mesurés sur cet échantillon témoin, pendant la même période de temps. La randomisation a entraîné une augmentation faible mais statistiquement significative du nombre d’annonces par page sur le réseau principal. Cette faible valeur s’explique par la faible performance du système.

Conclusion : avec ce genre d’approche, nous avons collectionné des données en intégrant peu d’aléatoire dans l’expérimentation tout en anticipant le prix d’une exploration plus importante. Il est possible de réaliser une exploration conditionnelle : par exemple si un utilisateur fait une requête commerciale, il est possible de réduire la collecte d’exploration. Pour répondre à plusieurs questions, si l’éditeur utilise un A/B test, il devra réaliser autant d’A/B test qu’il possède d’hypothèses. Avec une approche causale, il peut répondre aux questions contrefactuelles via une simple collecte optimisée. Cependant, avec la tendance du *Big Data*, les modèles statistiques doivent intégrer des milliers de paramètres. Cette explosion du nombre de paramètres entraîne des problèmes de variance.

5 Réduction de la variance

Dans cette section, deux approches orthogonales seront présentées. La première est rendue possible par la structure du graphe causal présenté dans la

Figure 5. La deuxième est plus générale et plus simple à mettre en place.

5.1 Réduction de la variance par ré-équilibrage

Nous nous intéressons cette fois au nombre moyen de clics (y) réalisés par des utilisateurs soumis à des sélections s . Les utilisateurs cliquent sans connaître les probabilités estimées de clics q et les prix d'enchères b des publicités leurs sont affichées. Nous allons intervenir en faisant varier les probabilités de clics q (c'est à dire que nous allons changer le modèle). En faisant varier q nous allons observer l'évolution de Y .

Les poids selon la sélection s sont :

$$w(\omega) = \frac{\mathbb{P}^*[\omega_i]}{\mathbb{P}[\omega_i]} = \frac{\mathbb{P}^*[s|x, a, b]}{\mathbb{P}[s|x, a, b]} \quad (16)$$

où

$$\mathbb{P}[s|x, a, b] = \int_q \mathbb{P}[s|x, a, b] \mathbb{P}[q|a, b] \quad (17)$$

$$\mathbb{P}^*[s|x, a, b] = \int_q \mathbb{P}[s|x, a, b] \mathbb{P}^*[q|a, b] \quad (18)$$

Dans les équations (17) et (18), $\mathbb{P}[s|x, a, b]$ et $\mathbb{P}^*[s|x, a, b]$ n'incluent plus la probabilité de clic q puisque nous intégrons selon toutes les valeurs de q . Nous pouvons à nouveau estimer l'espérance de Y en utilisant les relations causales de la Figure 6 et une isolation :

$$\begin{aligned} Y^* &= \int_{\omega \in \Omega \setminus \Omega_R} y \mathbb{P}^*(u, v, x, a, b, s, y) \\ &= \int_{\omega \in \Omega \setminus \Omega_R} y \frac{\mathbb{P}^*[s|x, a, b]}{\mathbb{P}[s|x, a, b]} \mathbb{P}^*(u, v, x, a, b, s, y) \\ &\approx \frac{1}{n} \sum_{i=1}^n y_i \frac{\mathbb{P}^*[s_i|x_i, a_i, b_i]}{\mathbb{P}[s_i|x_i, a_i, b_i]}. \end{aligned} \quad (19)$$

Cette approche absorbe la variance due à la sélection s montrée à l'utilisateur (représentée dans la Figure 10 avant ré-équilibrage et la Figure 11). Lorsque nous faisons fait varier la *mainline reserve* comme dans la section 4.3, nous constatons dans la Figure 12 que les intervalles de confiances sont réduits sur l'ensemble des variations ρ^* .

Il est alors possible, avec un système d'attribution des sélections s randomisées, d'estimer le nombre de clics moyen que nous aurions obtenue avec un modèle différent. Sous condition que l'éditeur regroupe les variables qui se trouvent entre le point d'intervention et le résultat qu'il veut moyenner, tout en produisant des bornes sur chaque chemin (nécessaire pour l'intégration), nous obtenons une réduction significative de la variance.

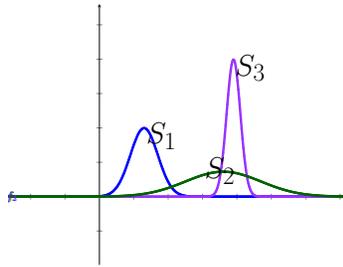


FIGURE 10 – Probabilités de clics q observées selon différentes sélections s

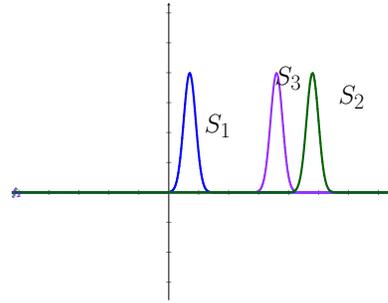


FIGURE 11 – Probabilités de clics q contrefactuelles selon différentes sélections s

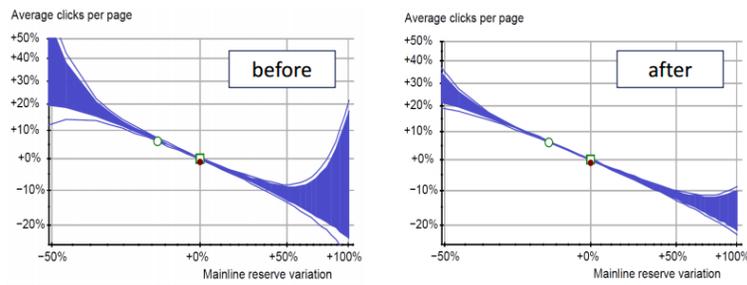


FIGURE 12 – Nombre moyen de clics contrefactuel avant/après ré-équilibrage

5.2 Réduction de la variance par une fonction de prédiction

Cette approche proposée par (Dudik, Dumitru, Langford, & Li, 2012) est une des plus utilisées. Elle repose sur une fonction de prédiction vérifiant que la variable de mesure se trouve dans un intervalle de confiance, calculé par une fonction de prédiction prenant en entrée les variables sur le chemin causal. Cette fonction de prédiction a notamment déjà été utilisée pour améliorer l'apprentissage de taux de clics dans (Claeys, Gancarski, Maumy-Bertrand, & Wassner, 2018). Voici un exemple illustratif : si nous regardons le nombre de clics moyen en fonction du temps, nous avons observé un comportement périodique . En pratique, sur certains sites, on constate par exemple que les gens cliquent plutôt en journée qu'en soirée. Cela a pour conséquence d'augmenter la variance. On peut alors rajouter au modèle causal une fonction ζ connue qui est la fonction de prédiction. L'idée est qu'elle serve à donner un nombre moyen de clics *a priori*. Nous ne traiterons pas de l'entraînement de cette fonction, le lecteur peut s'appuyer sur des modèles de régression usuelle. Une fois cette fonction construite, l'espérance de clic est partagée en deux parties : l'espérance du prédicteur et l'espérance estimée. Lorsqu'on étudie la distribution contrefactuelle, on peut facilement calculer ζ^* (puisque'elle est connue), et l'espérance contrefactuelle est calculée via les poids $w(\omega_i)$ (cf section 4.1). La Figure 13 représente le schéma causal de la Figure 6 en incluant la fonction ζ . Cette fonction intervient dans l'estimation de Y^* contrefactuelle, L'idée sous-jacente à cette méthode est que si l'erreur du prédicteur ($l(\omega_i) - \zeta(\omega_i)$) est faible, les variables ω_i auront un poids plus important et permettront de réduire la variance. Pour une variable de mesure $l(\omega)$, Y^* est définie par :

$$Y^* = \int_{\omega} l(\omega) \mathbb{P}^*[\omega] d\omega = \int_{\omega} \zeta(\omega) \mathbb{P}^*[\omega] d\omega + \int_{\omega} (l(\omega) - \zeta(\omega)) \mathbb{P}^*[\omega] d\omega \quad (20)$$

$$\approx \frac{1}{n} \sum_{i=1}^n \zeta^*(\omega_i) + \frac{1}{n} \sum_{i=1}^n (y_i - \zeta(\omega_i)) w(\omega_i) \quad (21)$$

Lorsque nous nous intéressons au clic moyen des utilisateurs y , nous définissons $\zeta(\omega_i) = \zeta(x_i, q_i, s_i)$ et nous obtenons :

$$Y^* = \frac{1}{n} \zeta^*(x_i, q_i, s_i) + \frac{1}{n} \sum_{i=1}^n (y_i - \zeta(x_i, q_i, s_i)) w(\omega_i)$$

L'estimateur Y^* est doublement robuste :

- si la fonction de prédiction est correcte, le premier terme $\frac{1}{n} \zeta^*(\omega_i)$ donnera l'estimation sans erreur (puisque $l(\omega_i) - \zeta(\omega_i) \approx 0$).
- si la fonction de prédiction n'est pas performante (beaucoup d'erreurs), les poids $w(\omega)$ viendront corriger l'erreur d'estimation.

Cette double robustesse aboutit à une réduction de la variance quelle que soit la fonction ζ .

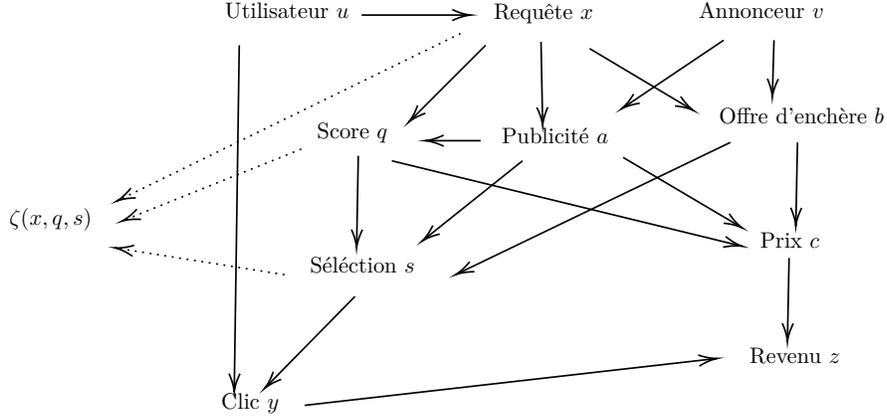


FIGURE 13 – Ajout de la fonction ζ dans le réseau Bayésien

6 Approche contrefactuelle pour l'apprentissage d'un modèle optimal

Nous souhaitons répondre à la question contrefactuelle suivante : «Que se serait-il passé si, pendant la collecte de données, l'éditeur avait développé un modèle M_1 de probabilité de clics q , plutôt qu'un modèle M_2 ?». Ce dernier peut alors comparer les deux modèles via la différence $\Delta = q_{M_1} - q_{M_2}$ et déterminer le meilleur modèle. Les deux modèles M_1 et M_2 étant très proches, la fonction de prédiction ζ permet de limiter les effets externes. Notons θ les paramètres de la variable d'intervention. Par exemple, nous voulons savoir quelle aurait été la performance du moteur de placement publicitaire si nous avions utilisé différentes valeurs θ pour modéliser la probabilité de clics q . Soit $\mathbb{P}^\theta[\omega]$ la factorisation de Markov contrefactuelle associée à cette intervention (par exemple celle de l'équation (5)). Soit Y^θ l'espérance contrefactuelle de $l(\omega)$ sous la distribution P^θ . Les données d'apprentissage sont collectées à partir d'une seule expérience associée à une valeur de paramètre initial θ^0 puis une nouvelle collecte utilisant θ^* est réalisée. L'objectif ici est de récolter des données en exploitant le paramètre θ optimal :

$$\theta^* = \operatorname{argmax}_{\theta} \hat{Y}^\theta.$$

Nous définissons l'estimation \hat{Y}^θ par :

$$\begin{aligned} \hat{Y}^\theta &= \int_{\omega} l(\omega) w_\theta(\omega) \mathbb{P}[\omega] \\ &\approx \frac{1}{n} \sum_{i=1}^n l(\omega_i) w_\theta(\omega_i) \end{aligned}$$

avec :

$$w_\theta(\omega_i) = \frac{\mathbb{P}^\theta[\omega_i]}{\mathbb{P}[\omega_i]}.$$

Nous pouvons calculer la dérivée partielle de Y^θ et de w_θ selon θ :

$$\begin{aligned} \frac{\partial Y^\theta}{\partial \theta} &= \int_{\omega} l(\omega) w'_\theta(\omega) \mathbb{P}[\omega] d\omega \\ &\approx \frac{1}{n} \sum_{i=1}^n l(\omega_i) w'_\theta(\omega_i). \end{aligned}$$

Lorsque nous nous intéressons aux probabilité de clics q nous obtenons :

$$w_\theta(\omega) = \frac{\mathbb{P}^\theta[q|x, a]}{\mathbb{P}[q|x, a]} \quad (22)$$

$$dw_\theta(\omega) = \frac{\partial w_\theta(\omega)}{\partial \theta} = w_\theta(\omega) \frac{\partial \log \mathbb{P}^\theta[q|x, a]}{\partial \theta} \quad (23)$$

Lorsque nous collectons les données avec un modèle θ_0 , nous procédons au calcul de la dérivée de q par rapport à θ en θ^0 .

Exemple : L'éditeur sélectionne des publicités a selon l'ordre décroissant associé à leur valeur $b * q^\alpha$ avec un $0 < \alpha < 1$ (Lahaie & McAfee, 2011). L'idée est de baisser la valeur de q par rapport à b , ce qui peut être efficace lorsque le modèle q fait beaucoup d'erreurs. L'éditeur définit $k \in \mathbb{N}$ groupes, et leur affecte une valeur α_k et une *mainline reserve*. La Figure 23 montre pour un seul groupe,

- En abscisse, la *mainline reserve*.
- En ordonnée, la valeur de α .
- Une valeur en noir qui est la valeur b de l'enchère (unités arbitraires allant de 164 à 169). Cette valeur correspond à une estimation du surplus de l'annonceur. La courbe noire montre l'évolution de cette valeur en fonction de α et du nombre de places disponible. Plus le nombre de places disponibles est grand, plus la valeur d'une place est basse. Ce sont ces courbes qui correspondent aux lignes de niveau (contrefactuelles)
- Un point de départ en rouge : Un nombre de publicités qui seront affichées (non renseigné car arbitraire) sur le résultat de la recherche. Cette valeur correspond au nombre moyen de publicités qui apparaîtront au dessus des résultats du moteur de recherche (c'est à dire dans la *mainline*). La courbe rouge décrit l'évolution de ce nombre en fonction de α et une *Mainline reserve*. L'évolution de cette valeur par rapport à son point d'origine est renseignée (allant de -6% à +10%). Cette courbe est nécessairement corrélée au nombre de places qui sont proposées lors de l'enchère.

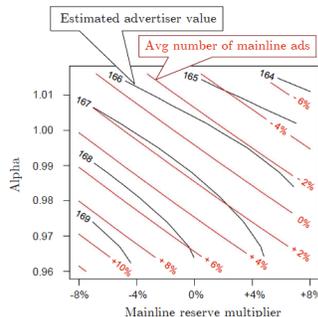


FIGURE 14 – Évolution de la valeur moyenne d’une place par les publicitaires, et du nombre de places achetées, en fonction de α et du nombre de place disponibles

L’idée ici est de montrer comment le *publisher* peut optimiser le α et la *mainline reserve* séparément pour chaque groupe. On procédant ainsi pour chaque groupe, l’éditeur peut choisir une *mainline reserve* qui maximise sa valeur pour tous les annonceurs, sans que le nombre moyen de publicités dans la *mainline* excède un seuil prédéfini.

Le lecteur peut trouver des détails sur le calcul des intervalles de confiance autour de \hat{Y}^θ dans (Bottou et al., 2013). Une approche séquentielle est également possible, nous optimisons les paramètres évalués en utilisant un modèle de bandit (par exemple Thompson Sampling). Les données d’apprentissage sont régulièrement collectées (voir Figure 4) mais cette fois le système anticipe le changement de la distribution. Nous parlons alors d’**apprentissage externalisé**. En effet, le paramètre θ est extérieur au système et impacte le score q . Cette action introduit une notion d’équilibre.

7 Notion d’équilibre

Le système causal décrit dans la section 2 inclut un indicateur a qui est nécessaire à l’affichage d’une publicité. Un seuil de a est fourni par un expert et renseigne sur la qualité de la publicité (par exemple : commerciale/non commerciale). Ce seuil peut se voir également comme une variable quantitative. Si nous augmentons ce seuil, les utilisateurs verront moins de publicités, mais elles seront de meilleure qualité. De plus, une augmentation pourra faire monter les enchères pour les places disponibles. En effet, les utilisateurs seront plus enclins à cliquer.

L’éditeur souhaite répondre à la question suivante : «quelles seraient les performances du système s’il changeait cette valeur qui se compensait à long terme avec l’augmentation de la valeur des enchères?». Cette compensation entre le nombre d’enchères disponibles et leurs valeurs peut se voir comme une notion d’équilibre. Nous considérons par la suite que cet équilibre est constant dans le temps. Pour des paramètres constants, l’équilibre consiste à dire que

la relation entre ces deux paramètres est linéaire. La causalité est ici dans les deux sens. Nous nous intéressons aux offres d'enchères b , au revenu généré par l'utilisateur z après son action y et au paramètre θ . Les variables observées ω sont des fonctions linéaires de θ et b :

$$d\omega = \frac{\partial\omega}{\partial\theta} + \sum_a^n \frac{\partial\omega}{\partial b_a} db_a$$

Un modèle g est fixé pour représenter l'impact des revenus z associés aux actions d'utilisateurs y et à la publicité a : $b_a = g(y_a, z_a)$:

$$db_a = \frac{\partial b_a}{\partial y_a} dy_a + \frac{\partial b_a}{\partial z_a}$$

Ici $d\omega$ et db_a représente un sous espace tangent. Nous définissons une valeur V_a constante telle que :

$$\begin{aligned} V_a &= \frac{\partial Z_a}{\partial Y_a} \\ &\approx \frac{\frac{\partial Z_a}{\partial b_a}}{\frac{\partial Y_a}{\partial b_a}} = \frac{\frac{\partial \mathbb{E}_{b,\theta}[z_a]}{\partial b_a}}{\frac{\partial \mathbb{E}_{b,\theta}[z_a]}{\partial y_a}}. \end{aligned}$$

Nous cherchons l'intersection de ces deux espaces (voir Figure X), représentant l'équilibre :

$$\forall a' \in \mathcal{A}, \left(V_{a'} \frac{\partial^2 Y_{a'}}{\partial b_{a'} \partial \theta} + \frac{\partial^2 Z_{a'}}{\partial b_{a'} \partial \theta} \right) d\theta + \sum_{a \in \mathcal{A}} \left(V_{a'} \frac{\partial^2 Y_{a'}}{\partial b_{a'} \partial b_a} + \frac{\partial^2 Z_{a'}}{\partial b_{a'} \partial b_a} \right) db_a = 0 \quad (24)$$

Les interventions dépendent ici d'un paramètre continu θ . Par exemple, lorsque nous souhaitons savoir quelles auraient été les performances du moteur de placement publicitaire si nous avions utilisé un modèle de notation paramétré selon θ . Nous rappelons que $\mathbb{P}^\theta[\omega]$ représentant la factorisation Markovienne contrefactuelle associée à cette intervention. Soit Y^θ l'espérance contrefactuelle de $l(\omega)$ sous la distribution \mathbb{P}^θ . La dérivée présentée de l'équation (20) donne

$$\begin{aligned} \frac{\partial Y^\theta}{\partial \theta} &= \int_\omega [l(\omega) - \zeta(v)] w'_\theta(\omega) \mathbb{P}[\omega] d\omega \\ &\approx \frac{1}{n} \sum_{i=1}^n [l(\omega_i) - \zeta(v_i)] w'_\theta(\omega_i). \end{aligned}$$

où $w_\theta(\omega)$ et $w'_\theta(\omega)$ calculé dans l'équation (23). Le remplacement des expressions $\mathbb{P}[\omega]$ et $\mathbb{P}^\theta[\omega]$ par les factorisations de Markov correspondantes offre de nombreuses possibilités de simplifier le ratio de repondération $w'(\omega)$. Le terme $w(\omega)$ est calculé via l'équation (7). La dérivée de $\log \mathbb{P}^\theta[\omega]$ dépend uniquement des facteurs paramétrés par θ . Par conséquent, afin d'évaluer $w'(\omega)$ nous avons

seulement besoin de connaître les quelques facteurs affectés par la fonction d'intervention. Les dérivées d'ordre supérieur peuvent être estimées à l'aide de la même approche. Par exemple, nous avons :

$$\begin{aligned}\frac{\partial^2 Y^\theta}{\partial \theta_i \partial \theta_j} &= \int_{\omega} [l(\omega) - \zeta(v)] w''_{ij}(\omega) \mathbb{P}[\omega] d\omega \\ &\approx \frac{1}{n} \sum_{i=1}^n [l(\omega_i) - \zeta(v_i)] w''_{ij}(\omega_i)\end{aligned}$$

avec

$$\begin{aligned}w''_{ij}(\omega_i) &= \frac{\partial^2 w_\theta(\omega)}{\partial \theta_i \partial \theta_j} = w_\theta(\omega) \frac{\partial \log \mathbb{P}^\theta(\omega)}{\partial \theta_i \partial \theta_j} \\ &= w_\theta(\omega) \frac{\partial \log \mathbb{P}^\theta(\omega)}{\partial \theta_i} \frac{\partial \log \mathbb{P}^\theta(\omega)}{\partial \theta_j} + w_\theta(\omega) \frac{\partial^2 \log \mathbb{P}^\theta(\omega)}{\partial \theta_i \partial \theta_j}\end{aligned}$$

La valeur d'équilibre dans l'équation (24) peut ensuite être utilisée pour estimer la valeur contrefactuelle de Y lorsque l'éditeur change son paramètre θ et que les annonceurs dans l'ensemble \mathcal{A} ajustent leurs offres b en réponse :

$$dY = \left(\frac{\partial Y}{\partial \theta} + \sum_{a \in \mathcal{A}} \frac{\partial Y}{\partial b_a} \frac{\partial b_a}{\partial \theta} \right) d\theta.$$

Conclusion : nous avons intégré un modèle causal fonctionnant par pondération afin de répondre à des problématiques de publicité online. Nous obtenons avec une telle approche des estimations fiables sur un taux de clic immédiat et un système viable à long terme. La collecte, liée à l'exploration, peut être coûteuse, il est donc intéressant de savoir la quantité à randomiser. Pour ne pas être trop coûteuse, le temps qui lui est consacré doit être minimal et permettre de déduire des améliorations futures sur le système. Cependant, il y a des questions qui n'ont pas de réponse. L'exploration ne garantit pas la réponse à une question qui ne peut se vérifier par les données dont on dispose. Enfin, il est également nécessaire de disposer d'un historique détaillé et complet (pour toutes les pages et sur toutes les décisions). Il faut également veiller à ce que la collecte distingue les utilisateurs et les publicitaires, de façon à obtenir des échantillons indépendants.

8 Intuition causale

Cette section traite de la découverte de relation causale à partir d'une collecte de données. Elle présente via une deuxième étape, des indices sur les relation causale en évitant l'affichage randomisée dans l'expérience.

En effet, la première étape consistait à procéder à une expérience randomisée et donnait des observations sans facteur de confusion. Les facteurs de confusion donnent des causes communes sur des facteurs d'intervention. Si l'intervention

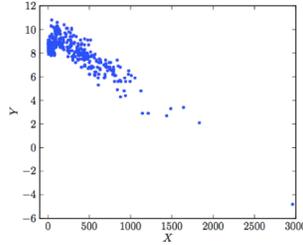


FIGURE 15 – Observation d’une variable Y par rapport à X

est aléatoire, il ne peut pas y avoir de facteur de confusion. La deuxième étape est l’étude d’intervention qui, sous supposition d’intervention, est équivalente à de l’expérience randomisée. Par exemple pour une position fixée, nous regardons différents paramètres. Mais comment déterminer les variables causales? Une observation visuelle des données peut aboutir à des *clusters* visibles qui sous entendent une variable causale non observée. Par exemple dans la Figure 15, il y a une corrélation positive et une variable cachée liée à la nature des *clusters*. Nous pouvons alors formuler les hypothèses suivantes : pour deux variables X et Y ,

1. X cause Y .
2. Y cause X .
3. X et Y ont des causes communes non observées.
4. X et Y sont indépendants.

Nous nous consacrons aux hypothèses 1 et 2. Nous observons pour différentes valeurs de $Y = \alpha X + \beta + \epsilon$ ou $\alpha, \beta \in \mathbb{R}$ et ϵ un bruit gaussien. Si X varie, c’est que l’hypothèse 2 est vérifiée. La détection du sens de causalité peut utiliser plusieurs techniques notamment :

- une fonction génératrice des moments :

$$F_{r,s} = \frac{1}{n} \sum_{i=1}^n X_i^r Y_i^s$$

pour r et s fixés.

- un espace de Hilbert

$$F = \frac{1}{n} \sum_{i=1}^n \phi(X_i, Y_i) \in \mathcal{H}$$

avec $\langle \phi() \phi() \rangle_K = K(., .)$.

- Un classifieur

$$F_w = \frac{1}{n} \sum_{i=1}^n \phi_w(X_i, Y_i).$$

Les classifieurs, notamment grâce aux réseaux de neurones, sont aujourd'hui les plus populaires. Chaque point est structuré par une fonction non linéaire et nous récoltons ces fonctions pour chaque point. Nous moyennons ensuite sur tous les exemples et nous appliquons à un classifieur.

Pour les données d'apprentissage, nous tirons des distributions symétriques pour générer des données artificiellement :

$$Y = f(X) + v(X)\epsilon.$$

Nous procédons aux étapes suivantes :

1. Nous choisissons un paramètre j .
2. Nous tirons $k \sim \mathcal{U}\{1, 2, \dots, j\}$ et $r, s \sim \mathcal{U}[0, J]$. Nous simulons un modèle de mélange gaussien de $\mu \sim \mathcal{N}(0, r)$ et $\sigma \sim \mathcal{N}(0, r)$.
3. Nous modélisons $f(X)$ et $v(X)$ à travers une fonction polynomiale cubique.
4. Nous ajoutons un bruit gaussien ϵ avec une variance aléatoire tirée uniformément dans $\mathcal{U}[0, J]$.
5. Nous calculons $Y_j = f(X_i) + v(X_j)\epsilon_j$ et on représente un nuage de points $X \leftarrow Y$ (voir Figure [15](#))
6. Nous représentons un nuage de points $X \rightarrow Y$ mais avec X et Y centrées réduites.
7. Nous répétons les étapes 5 et 6 mais cette fois avec la configuration $Y \rightarrow X$
8. Nous labelisons les points $\{X_i, Y_i\}$ comme appartenant à la classe #1 (X cause Y)
9. Nous labelisons les points $\{Y_i, Y_i\}$ comme appartenant à la classe #0 (Y cause X)

La génération de 100 000 points suffit à entraîner un réseau de neurones (79% des relations causales ont été correctement identifiées). Des expériences récentes identifient maintenant des relations causales à travers des images. Le lecteur peut trouver des informations complémentaires dans [\(Bottou et al., 2013\)](#).

9 Conclusion

Dans un premier temps, nous avons montré comment le raisonnement contre-factuel et l'inférence causale permettent à l'ingénieur de prévoir les conséquences des changements appliqués à un moteur de placement de publicités. Ces méthodes fonctionnent lorsque l'on introduit un peu d'aléatoire dans le moteur de placement lui-même. Cela ressemble donc à des expériences dites "randomisées" et se rapproche des techniques usuelles en statistiques causales mais avec une échelle bien plus grande. En revanche la section 8 présente une hypothèse inhabituelle selon laquelle les moments élevés de la distribution jointe observée des variables donnent des indices sur leur relation causale sans même avoir besoin d'introduire un aspect aléatoire dans l'expérience.

Références

- Audibert, J.-Y., Munos, R., & Szepesvari, C. (2007). Tuning bandit algorithms in stochastic environments. In *Algorithmic Learning Theory* (p. 150-165). Sendai, Japan. Consulté sur <https://hal.inria.fr/inria-00203487>
- Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002, mai). Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.*, 47(2-3), 235–256. Consulté sur <https://doi.org/10.1023/A:1013689704352> doi: 10.1023/A:1013689704352
- Blake, T., Nosko, C., & Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness : A large-scale field experiment. *Econometrica*, 83(1), 155-174. doi: 10.3982/ECTA12423
- Bottou, L., Peters, J., nonero Candela, J. Q., Charles, D. X., Chickering, D. M., Portugaly, E., ... Snelson, E. (2013). Counterfactual reasoning and learning systems : The example of computational advertising. *Journal of Machine Learning Research*, 14, 3207-3260. Consulté sur <http://jmlr.org/papers/v14/bottou13a.html>
- Chiffres-clés de google.* (s. d.). <https://www.webrankinfo.com/dossiers/google/chiffres-cles>. (Accessed : 2017-09-22)
- Claeys, E., Gancarski, P., Maumy-Bertrand, M., & Wassner, H. (2017). Regression tree for bandits models in a/b testing. In *Advances in intelligent data analysis xvi* (pp. 52–62). Cham : Springer International Publishing.
- Claeys, E., Gancarski, P., Maumy-Bertrand, M., & Wassner, H. (2018, septembre). *Dynamic allocation optimization in a/b tests using classification-based preprocessing.* Consulté sur <https://hal.archives-ouvertes.fr/hal-01874969> (working paper or preprint)
- Dudik, M., Dumitru, E., Langford, J., & Li, L. (2012, Oct). Sample-efficient Nonstationary Policy Evaluation for Contextual Bandits. *arXiv e-prints*, arXiv :1210.4862.
- Edelman, B., Ostrovsky, M., & Schwarz, M. (2007, March). Internet advertising and the generalized second-price auction : Selling billions of dollars worth of keywords. *American Economic Review*, 97(1), 242-259. Consulté sur <http://www.aeaweb.org/articles?id=10.1257/aer.97.1.242>
- Lahaie, S., & McAfee, R. P. (2011). Efficient ranking in sponsored search. *7th International Workshop on Internet and Network Economics (WINE 2011)*, 254–265, publisher = Springer.
- Lecoutre, B. (2004). Expérimentation, inférence statistique et analyse causale. *Intellectica*, 38(1), 193–245. Consulté sur https://www.persee.fr/doc/intel_0769-4113_2004_num_38_1_1713 doi: 10.3406/intel.2004.1713
- Li, L., Chu, W., Langford, J., & Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. *CoRR*, abs/1003.0146. Consulté sur <http://arxiv.org/abs/1003.0146>
- Milgrom, P. (2004). *Putting auction theory to work.* Cambridge University Press. doi: 10.1017/CBO9780511813825
- Myerson, R. B. (1981, février). Optimal auction design. *Math. Oper. Res.*,

- 6(1), 58–73. Consulté sur <http://dx.doi.org/10.1287/moor.6.1.58>
doi: 10.1287/moor.6.1.58
- New data why people hate ads.* (s. d.). <https://www.vieodesign.com/blog/new-data-why-people-hate-ads>. (Accessed : 2018-09-06)
- Pearl, J. (2009). *Causality : Models, reasoning and inference* (2nd éd.). New York, NY, USA : Cambridge University Press.
- Pearl, J. (2012). The do-calculus revisited. *CoRR*, *abs/1210.4852*. Consulté sur <http://arxiv.org/abs/1210.4852>
- Riley, J., & Samuelson, W. F. (1981). Optimal auctions. *American Economic Review*, *71*(3), 381-92. Consulté sur <https://EconPapers.repec.org/RePEc:aea:aecrev:v:71:y:1981:i:3:p:381-92>
- Robins, J. (1987). A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *Journal of Chronic Diseases*, *40*, 139S - 161S. Consulté sur <http://www.sciencedirect.com/science/article/pii/S0021968187800188> doi: [https://doi.org/10.1016/S0021-9681\(87\)80018-8](https://doi.org/10.1016/S0021-9681(87)80018-8)
- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *13*(2), 238–241. Consulté sur <http://www.jstor.org/stable/2984065>
- Sutton, R. S., & Barto, A. G. (1988). Reinforcement learning : An introduction. *IEEE Transactions on Neural Networks*, *16*, 285-286.
- Tang, D., Agarwal, A., O'Brien, D., & Meyer, M. (2010). Overlapping experiment infrastructure : More, better, faster experimentation. In *Proceedings 16th conference on knowledge discovery and data mining* (pp. 17–26). Washington, DC.
- Varian, H. (2007). Position auctions. *International Journal of Industrial Organization*, *25*(6), 1163-1178. Consulté sur <https://EconPapers.repec.org/RePEc:eee:indorg:v:25:y:2007:i:6:p:1163-1178>
- von Wright, G. H. (1971). *Explanation and understanding*. Cornell University Press.
- Woodward, J. (2005). Making things happen. *Philosophical Review*, *114*(4), 545–547.