



Mon rosbif est biaisé

Serge Abiteboul¹ et Julia Stoyanovich²

Une discrimination systématique par un algorithme est appelée un *biais*. Les biais peuvent avoir différentes origines, et ils correspondent tous à des situations de la vie réelle. Pour illustrer cela, pensez à cuisiner du rosbif.

Pat décide de publier un site internet de recettes familiales en commençant par la plus populaire : le rosbif. La mère de Pat, sa sœur et son frère préparent tous le rosbif à peu près de la même manière : « comme le fait grand-mère ». Ils utilisent tous la même recette — le même algorithme. Leur choix de cuisiner le rosbif de cette manière montre une sorte de biais de *familiarité* — privilégiant les goûts et les expériences familiers par rapport aux nouveaux.

Lorsqu'elle sélectionne une recette de rosbif familière de manière biaisée, Pat ne fait rien de mal — elle suit simplement une tradition familiale agréable. Mais dans d'autres cas, le biais d'habitude peut être nocif : il peut conduire à une discrimination illégale dans l'embauche, le prêt ou l'octroi d'une libération conditionnelle. Aussi inoffensive soit-elle, la recette de rosbif de la famille de Pat nous permettra de discuter différents types de préjugés.

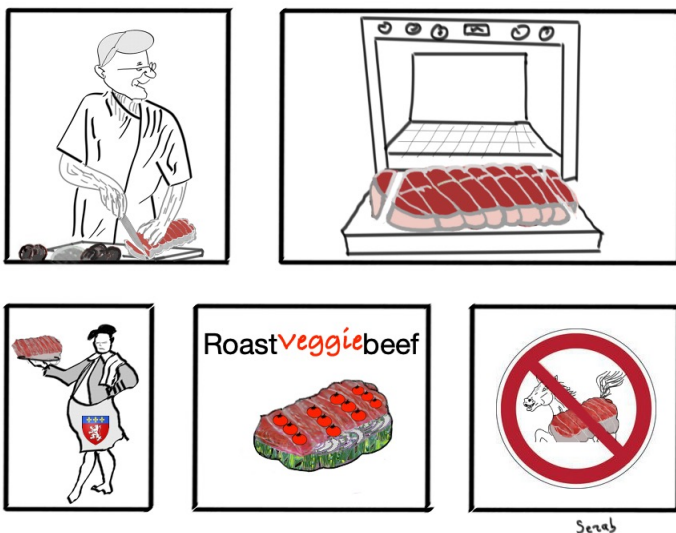
Comment choisir une recette de rosbif à suivre ? Commencez par sélectionner vos données d'entraînement, une recette « générique » que vous avez vu d'autres utiliser, et faites-la vôtre ! Dans le cas de Pat, le type de préjugé le plus visible vient des personnes dont elle adopte la recette : sa famille. Par exemple, au Maroc, la recette générique serait de préparer du bœuf avec du curcuma, du gingembre, du cumin et de la coriandre, ainsi que du jus de grenade. Par contre, en Alsace, la base du rosbif

1. Directeur de recherche Inria et membre de l'Académie des sciences.

2. Professeure à New York University.

n'est pas du tout du bœuf, mais de la viande de cheval (*ross* signifie cheval en alsacien). En clair, selon les personnes qu'on choisit, on obtient des recettes différentes. La famille de Pat est plutôt traditionnelle, ils utilisent du bœuf et le garnissent de persil, de sauge, de romarin et de thym (comme dans la chanson³). Le biais introduit par la sélection des données d'entraînement est parfois appelé biais *préexistant*, car il provient de recettes qui existaient avant que vous n'envisagiez même de faire vous-même ce rosbif.

Une pratique s'avère assez particulière dans la famille de Pat : ils découpent des tranches aux deux extrémités du rôti avant de le mettre au four. Pourquoi font-ils cela, demande Pat ? Un cousin propose : à cause de l'oxydation du rôti aux extrémités. Un autre n'est pas d'accord : parce que nous, les enfants, aimons notre rosbif bien cuit, et les petits morceaux sont alors hyper cuits. Pat demande à sa mère, sa tante et son oncle qui répondent : nous faisons ceci parce que maman faisait cela. Alors, quand elle en a l'occasion, elle demande à sa grand-mère qui répond : ma mère faisait comme ça. Mais pourquoi ? Parce que son four était trop petit pour contenir tout le rôti. C'est ce qu'on appelle le biais *technique*. En informatique, un tel biais peut être introduit, par exemple, parce qu'une représentation incorrecte des données a été choisie. Et ce n'est peut-être pas si facile à détecter car les programmeurs ont toujours tendance à supposer la perfection de leur code. Ou, du moins, ils remettent rarement en question un morceau de code qui fonctionne.



3. https://fr.wikipedia.org/wiki/Parsley,_Sage,_Rosemary_and_Thyme.

Pour introduire le dernier type de biais, revenons à l'Angleterre médiévale⁴. Il était une fois une ville avec deux chefs qui avaient inventé des recettes de rosbif concurrentes. Le duc qui avait entendu parler de ces deux nouvelles recettes se déplaça pour venir les déguster. Malheureusement, l'un des deux chefs eut un rhume ce jour-là. Ainsi, seul Maître Chiquart, un chef français originaire de Lyon, put faire déguster son rosbif. Le duc l'apprécia tant qu'il l'introduisit à la cour du roi. Ce fut le début de la saga de la recette du rosbif que l'on connaît aujourd'hui, alors que la recette de l'autre chef se perdit dans l'histoire, jusqu'au nom de ce chef. Sans ce rhume ce jour fatidique, la recette de rosbif de Pat aurait peut-être été très différente. C'est ce qu'on appelle le biais *émergent* : plus les gens utilisent la recette de Chiquart, plus elle deviendra populaire et plus les gens l'utiliseront. Cela rendra la préparation du rosbif de plus en plus uniforme.

Le site web de Pat n'est pas très populaire auprès des *millenials*, la génération Y ; beaucoup d'entre eux sont végétariens et se soucient peu de recettes de rosbif. Un biais *générationnel* ?

4. Attention : cette histoire est une pure fiction. Il n'y a aucune preuve réelle que le rosbif ait été inventé par un chef français, même si cela est fort probable.