



Comment coproduisons-nous notre environnement numérique marchand ?

Jean-Sébastien Vayre

► **To cite this version:**

Jean-Sébastien Vayre. Comment coproduisons-nous notre environnement numérique marchand ?. Les objets numériques création et consommation, Aug 2014, Ax-les-Thermes, France. pp.1-21, 2014. <hal-01056152>

HAL Id: hal-01056152

<https://hal.archives-ouvertes.fr/hal-01056152>

Submitted on 16 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comment coproduisons-nous notre environnement numérique marchand ?

Jean-Sébastien VAYRE

Centre d'Etude et de Recherche Travail Organisation Pouvoir (CERTOP – UMR 5044)

Université Toulouse II-Le Mirail

n° 19 (Sociologie, démographie)

Maison de la Recherche, bâtiment B26

5 allées Antonio Machado – 31058 Toulouse Cédex 9

Bureau C437

05.61.50.39.77

jsvayre@univ-tlse2.fr

<http://www.certop.cnrs.fr/VAYRE-Jean-Sebastien>

MOTS-CLES :

Big data, numérisation du commerce, apprentissage artificiel, personnalisation, coproduction

RESUME :

Avec le développement des Technologies de l'Information et de la Communication (TIC) au sein des diverses sphères de la société, nous sommes actuellement conduit à produire une grande quantité de données numériques (i.e. : les big data) durant nos activités quotidiennes. Ces big data constituent autant d'indices que les e-commerçants mobilisent afin de personnaliser automatiquement les environnements numériques de leurs consommateurs. Pour ce faire, ils développent différents systèmes de filtrage destinés à améliorer la qualité de leurs services. C'est pourquoi, afin de mieux comprendre ce processus, nous proposons dans cet article de rendre compte de la manière dont nous coproduisons nos environnements numériques marchands. Nous verrons ainsi que cette coproduction recouvre finalement une asymétrie de conception qui la distingue très nettement d'une cocréation.

INTRODUCTION



Capture d'écran 1 : Exemple de recommandations marchandes

Comment le système de recommandation de *Teddy Smith* a-t-il pu faire l'hypothèse que les quatre produits ci-dessus pourraient nous intéresser ?

Durant l'ensemble de nos activités quotidiennes, nous sommes conduits à mobiliser de façon toujours plus intensive les Technologies de l'Information et de la Communication (TIC). Par voie de conséquence, nous sommes également amené à produire une très grande quantité de données numériques (i.e. : les big data) qui sont autant d'indices renseignant nos comportements de consommation (Kessous, 2012 ; Vayre, 2014). Et, c'est précisément à partir de ces données que le système de recommandation de *Teddy Smith* a pu réaliser un certain nombre d'inférences sur nos préférences de façon à nous communiquer les quatre recommandations de la capture d'écran 1.

En outre, avec le développement des big data, ces différents systèmes de filtrage que recouvrent les outils de personnalisation des environnements numériques marchands sont à la fois plus répandus et plus intelligents. Aussi, bien que les chercheurs en sciences humaines et sociales aient déjà cherché à rendre compte des hypothèses et des théories générales sous-tendus par ces dispositifs (e.g. : Baiocchi et Forest, 2014 ; Ménard, 2014) ces derniers semblent avoir fait l'impasse sur l'apprentissage artificielle qui est au cœur de ces technologies. Or, les hypothèses et les théories sous-tendues par ces dispositifs sont la résultante d'un apprentissage effectué par un ou plusieurs agent(s) artificiel(s).

Partant, à partir d'une ethnographie multi-située (Marcus, 1995) composée de plusieurs matériaux d'enquêtes¹, nous proposons dans cet article de rendre compte de la manière dont nous coproduisons nos environnements numériques marchands. Comment cette nouvelle

¹ Cet article repose en effet sur une enquête multi-située (Marcus, 1995) consistant à suivre l'ensemble du processus de traduction permettant aux interactions Hommes-Machines de devenir des données (non-)structurées, des informations puis des documentations (cf. Vayre, 2014). Cette étude réunie pour le moment l'ensemble des matériaux suivants : 1 ouvrage de référence dans le domaine de l'apprentissage artificiel (Cornuéjols et Miclet, 2010), 147 documents d'archives du *New York Times*, 15 entretiens semi-directifs conduits auprès de professionnels des données (i.e. : *data-scientists*), 20 présentations d'offres proposant des technologies marchandes big data, 11 retours d'expériences d'entreprises ayant adoptés des technologies marchandes big data, et 2 observations participantes de 6 mois effectuées dans deux start-up toulousaines développant des technologies marchandes big data. Aussi, bien que dans ces travaux nous ne référons pas directement l'ensemble de nos données de terrain et que nous avons mobilisé des exemples fictifs pour des raisons pédagogiques (et parfois de propriété intellectuelle), cet article n'en reste pas moins une synthèse construite à partir d'un ensemble de données empiriques.

matière première que composent les big data est-elle produite au quotidien (section 1) ? Comment les agents artificiels peuvent-ils apprendre des connaissances à partir des données numériques que constituent les big data (section 2) ? Quelles sont les grandes techniques d'apprentissage artificiel et, du point de vue des agents artificiels, quels sont les types de connaissances que ces techniques permettent de développer (section 3) ?

Dans la première section, nous verrons que la nouvelle matière première que constituent les big data est née des évolutions sociotechniques réalisées dans le domaine des TIC. Ces dernières recouvrent en effet un gigantesque système de production et de stockage des traces d'usages que sont les big data. Nous montrons alors comment les marchands² peuvent concevoir des agents artificiels capables d'inférer des connaissances à partir des big data de façon à personnaliser automatiquement les environnements numériques de leurs consommateurs. Nous verrons ainsi que l'apprentissage d'un agent artificiel fait nécessairement l'objet d'un problème et d'un scénario plus ou moins détaillé (section 2). Et, nous présentons les différentes techniques d'apprentissage et les différents types de connaissances que les agents artificiels peuvent développer (section 3). En conclusion, nous soutenons que, du point de vue du marché, la relation symbiotique qu'instituent les agents artificiels sous-tend une certaine asymétrie technique, cognitive et économique qui est bien la résultante d'un processus de coproduction qui ne doit cependant pas être confondu avec un processus de cocréation.

1 La captation des traces d'usages

Comme nous l'avons exposé en introduction, ce que les professionnels du marché nomment aujourd'hui les big data compose une sorte de nouvelle matière première destinée à alimenter les systèmes de personnalisation des environnements numériques marchands. En d'autres termes, les big data peuvent être considérées comme les éléments fondamentaux des nouvelles formes de coproduction des services proposés par les e-commerçants.

Partant, comment cette nouvelle matière première que composent les big data est-elle produite au quotidien ?

Dans la section 1.1, nous verrons que le développement des big data est directement associé aux évolutions réalisées dans le domaine des TIC puisque ces dernières recouvrent une sorte d'immense système de production et de captation des big data. Ensuite, nous verrons que ces big data sont finalement des traces d'usage qui constituent avant tout des indices permettant aux marchands de mieux connaître nos pratiques et nos préférences de consommation. En ce sens, nous soulignons les principales implications épistémologiques liées aux usages des big data (section 1.2).

1.1 Le système de production et de captation des big data

La nouvelle matière première que constituent les big data est née des évolutions sociotechniques réalisées dans le domaine des TIC (Vayre, 2014).

En effet, comme le montre Dominique Boullier (2004), les TIC constituent des médiations qui peuvent être comprises à travers un double processus d'appropriation-captation : elles permettent aux consommateurs de s'approprier des univers plus ou moins étrangers dans le même temps qu'elles permettent aux offreurs de pénétrer le monde des usagers³. Avec le

² Précisons qu'à travers l'expression « les marchands » nous souhaitons désigner le couple que forment les e-commerçants et les concepteurs des technologies de personnalisation des environnements numériques marchands qui sont souvent issus d'une entreprise externe dotée d'une forte expertise dans le domaine de l'informatique.

³ C'est très précisément ce double potentiel d'action que désigne le concept d'habîète (Boullier, 2004).

développement des réseaux de télécommunication, des centres de données, de l'informatique en nuage, de l'apprentissage artificiel (i.e. : *machine learning*), des capteurs⁴ et des technologies de visualisation, les TIC composent un immense système de production et de captation des big data que les offreurs mobilisent pour mieux connaître les consommateurs et pour mieux les documenter (Cochoy, 2004 ; Vayre, 2014).

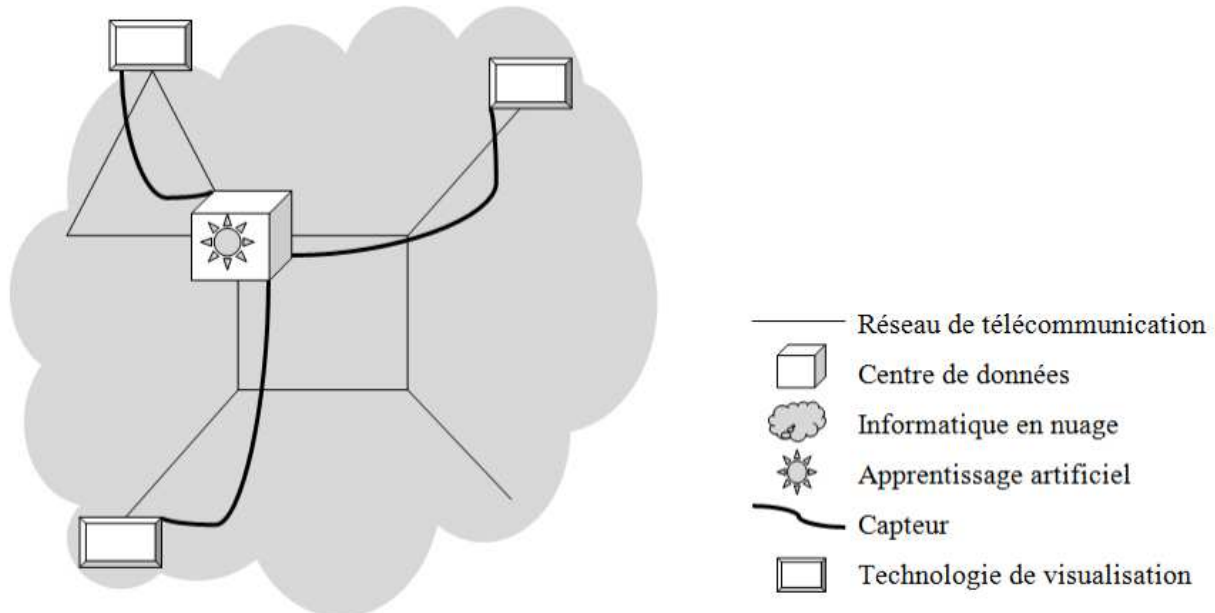


Schéma 1 : Le système de production et de captation des big data

C'est donc par le biais du système de production et de captation des big data que recouvrent les TIC que chacun d'entre nous est conduit à produire quotidiennement une grande quantité de données numériques durant ses activités :

- de loisir à travers les usages, par exemple, de *Facebook* ;
- de travail à travers les usages, par exemple, de *LinkedIn* ;
- de consommation à travers les usages, par exemple, d'*Amazon*.

Et, ce sont ces données que les professionnels du marché nomment les big data.

1.2 Les big data comme traces d'usages

En sociologie, il existe plusieurs façons de qualifier les big data. Emmanuel Kessous (2012) parle par exemple de dépôts d'attention laissés par les utilisateurs durant leurs activités numériques. Ici, les big data sont alors comprises comme les éléments fondamentaux d'une nouvelle économie de l'attention puisque ces dépôts sont stockés et exploités à travers un processus d'économicisation bien singulier (Akrich, 1989 ; Kessous, 2012).

Mais d'une manière plus générale, dans le domaine des sciences humaines et sociales, les big data sont considérées comme des traces d'usages (Mille, 2013). D'un point de vue sociologique, il est alors assez classique de s'attacher à mettre à jour le procès de fabrication des big data de façon à souligner l'importance des contextes sociomatériels à l'intérieur desquels elles sont produites (Jeanneret, 2013 ; Leleu-Merviel, 2013). En synthèse, l'argument est que les données brutes n'existent pas, et donc, que cette sorte de matière

⁴ e.g. : les ordinateurs et les téléphones portables, les tablettes numériques, les montres GPS, etc. ; c'est-à-dire, d'une manière générale, l'ensemble des objets connectés.

première que composent les big data ne constitue en rien une ressource naturelle (Gitelman, 2013).

Ainsi, comme dans le cas des données scientifiques, les big data sont construites selon différentes conventions sociales et techniques de façon à pouvoir re-présenter un phénomène particulier (Desrosières, 2008 ; Latour, 1993). Cependant, à la différence des données scientifiques, les big data ne découlent pas, dans la plupart des cas, d'un dispositif méthodologique rigoureusement contrôlé. Car, les big data sont souvent utilisées selon des objectifs différents de ceux qui ont motivé leur captation. C'est pourquoi, les traces d'usages que recouvrent les big data sont généralement qualifiées de non-sollicitées dans le sens où elles ne sont pas toujours construites et recueillies selon les mêmes objectifs que ceux qui orientent leur utilisation effective (Vayre, 2013).

Dès lors, la définition qu'Yves Jeanneret (2011) propose de la notion de trace apparaît comme particulièrement intéressante. Selon lui, la trace peut être comprise à travers les quatre dimensions suivantes :

- l'empreinte (i.e. : la trace est la marque physique laissée sur les objets) ;
- l'indice (i.e. : la trace engage une explication causale orientée vers le passé) ;
- l'inscription (i.e. : la trace est un artefact médiatique) ;
- le tracé (i.e. : la trace relève du monde graphique et donne forme aux images, aux récits et aux symboles).

De ce fait, en soulignant par exemple que les avis des consommateurs sont d'abord de l'ordre de l'inscription, cette catégorisation permet de mieux rendre compte du caractère médiatique des big data, et donc, des différents biais qui peuvent être impliqués par les cadres de leur production. Autrement dit, à l'image du guide touristique et à travers, par exemple, le site Internet d'un e-commerçant, le système de captation que recouvre les TIC ne fait pas qu'enregistrer les usages : « [il] les façonne puisque son organisation documentaire fait programme et définit des styles de pratique » (Jeanneret, 2013, p. 52).

Ce point permet alors d'insister sur la nécessaire vigilance épistémologique dont doivent faire preuve tout ceux qui mobilisent les big data dans le but de construire des connaissances. Ajoutons que cette vigilance épistémologique est d'autant plus importante que, dans le domaine marchand, ces connaissances ont généralement pour finalité d'alimenter des outils d'aide à la décision⁵, et donc, d'orienter l'action.

2 Le scénario d'apprentissage

À travers le système de production et de captation des big data que composent les TIC, chacun d'entre nous est conduit à produire quotidiennement de très nombreuses traces d'usages qui sont autant d'indices renseignant nos comportements de consommation. Et, c'est très précisément à partir de ces traces que les professionnels du marché peuvent personnaliser automatiquement nos environnements numériques marchands.

Pour ce faire, comme nous venons de le souligner, la première étape, qui n'est pas toujours effectuée avec attention, consiste à contrôler la qualité épistémique des big data. La seconde, qui fait aujourd'hui l'objet d'un travail souvent considérable, est de développer des agents

⁵ Précisons que les systèmes de recommandation et plus largement les divers dispositifs de personnalisation des environnements numériques marchands sont ici considérés comme des outils d'aide à la décision du consommateur (Vayre, 2013).

artificiels capables d'inférer des connaissances à partir des big data de façon à pouvoir prendre automatiquement la décision d'informer tel ou tel consommateur sur tel ou tel bien et/ou service de consommation.

Dans cette section, nous souhaitons traiter le questionnement suivant : comment les agents artificiels peuvent-ils apprendre des connaissances à partir des données numériques que constituent les big data ?

Nous verrons alors que tout apprentissage artificiel commence par un problème d'apprentissage (section 2.1). Nous montrons ensuite que, paradoxalement, pour pouvoir développer des connaissances solides, l'apprentissage d'un agent artificiel doit nécessairement être biaisé (section 2.2). Et, finalement, nous dégagons les différents cadres d'interaction permettant à un agent artificiel d'apprendre des connaissances sur son environnement (section 2.3).

2.1 Le problème d'apprentissage

Du point de vue des sciences de l'artificiel, tout apprentissage commence par un problème d'apprentissage (Cornuéjols et Miclet, 2010). Aussi, afin de bien comprendre en quoi consiste un tel problème et avant d'exposer un exemple concret, nous souhaitons revenir sur trois notions fondamentales dans le domaine de l'apprentissage artificiel :

- l'espace de description \mathcal{X} permet de désigner l'ensemble des exemples x (i.e. : les données entrantes) ;
- l'espace des sorties \mathcal{Y} permet de désigner l'ensemble des actions y associées à chaque exemple x (i.e. : les données sortantes) ;
- l'espace des hypothèses \mathcal{H} permet de désigner l'ensemble des hypothèses h sous-tendues par l'échantillon de données noté $S = [(x_1, y_1), \dots, (x_n, y_n)]$.

D'une manière générale, une activité d'apprentissage artificiel consiste alors à rechercher, au sein de l'espace des hypothèses \mathcal{H} , une fonction cible h^* qui soit la plus cohérente possible avec l'ensemble de données entrantes x et sortantes y .

Afin de permettre l'apprentissage d'un agent artificiel, il est donc nécessaire de spécifier les formes de l'espace de description \mathcal{X} , de l'espace des sorties \mathcal{Y} et de l'espace des hypothèses \mathcal{H} . Car, l'espace de description \mathcal{X} et celui des sorties \mathcal{Y} peuvent être composés de différents descripteurs (i.e. : attributs/variables) plus ou moins homogènes⁶. De même, l'espace des hypothèses \mathcal{H} renvoient à un ensemble de connaissances qui peuvent être représentées par le biais de différents langages de description $\mathcal{L}_{\mathcal{H}}$ ⁷.

Le problème d'apprentissage d'un agent artificiel peut alors être présenté à travers deux étapes (Cornuéjols et Miclet, 2010).

La première consiste à déterminer le contexte d'action et la tâche que l'agent artificiel doit apprendre à réaliser. Par exemple, imaginons que nous soyons un e-commerçant proposant des vélos de route et que nous souhaitons développer un agent artificiel capable d'apprendre en temps réel le concept « beau vélo » en fonction des trois descripteurs binaires que sont :

⁶ De façon non-exhaustive, il s'agit, par exemple, des descripteurs binaires (e.g. : oui/non), nominales (e.g. : une couleur), numériques (e.g. : un chiffre), nominales arborescents (e.g. : une hiérarchie), numériques séquentiels (e.g. : une suite de chiffres), etc.

⁷ C'est pourquoi, par exemple, l'apprentissage de classification, de régression, de distribution de probabilité, de concept, de graphe, etc., ne renvoient pas aux mêmes types de langage $\mathcal{L}_{\mathcal{H}}$.

- « Cadre atypique » (C : 1-oui/0-non) ;
- « Bicolore » (B : 1-oui/0-non) ;
- « Vintage » (V : 1-oui/0-non).

Précisons alors que notre objectif est de réaliser cette apprentissage à partir des photographies suggérées par notre outil de recommandation : les vélos mis en avant sont considérés comme « beaux » ou « pas beaux » *si et seulement si* les visiteurs ont respectivement cliqué sur les mentions « j'aime » ou « je n'aime pas » situées en dessous de chaque suggestion :

- « Like » (L : 1-oui/0-non).

La deuxième étape consiste alors à trouver le langage de description $\mathcal{L}_{\mathcal{H}}$ approprié au contexte et à la tâche d'apprentissage. Dans notre exemple, il s'agirait de la programmation logique inductive et plus particulièrement des formes normales conjonctives où le symbole « ^ » est alors considéré comme un connecteur de conjonction. De ce fait, $(C = 1) \wedge (B = 0) \wedge (V = 0) \wedge (L = 0)$ constitue par exemple une hypothèse de l'espace \mathcal{H} .

Notons alors que notre exemple renvoie à un problème d'apprentissage bien spécifique : celui de l'apprentissage supervisé de concept par induction. Et, plus formellement, ce problème peut être exposé comme suit. L'environnement procure des données d'entrée notées x (cf. les attributs C , B et V). Et, à chacune de ces données correspond une donnée de sortie notée y (cf. l'attribut L) que l'oracle (cf. le groupe des visiteurs) a déterminé à partir d'une fonction f qu'il est seul à connaître. L'apprenant artificiel bénéficie alors d'un échantillon de couples (x_i, y_i) tel que $(x_i, y_i) = (x_i, f(x_i))$. Son problème consiste ainsi à deviner la fonction f de l'oracle, ou, à tout le moins, à trouver une approximation optimale h^* de cette fonction.

Autrement dit, dans notre exemple, le problème que nous posons à notre agent artificiel est, dans une certaine mesure, d'apprendre les préférences de nos visiteurs.

2.2 Le biais d'apprentissage

Dès lors, comment un apprenant artificiel s'y prend-t-il pour deviner la connaissance de l'oracle ?

Reprenons notre exemple et imaginons cette fois-ci que nous venons de mettre en ligne notre site de vente de vélo de route. Pour le moment, ne disposant que de peu de références et n'étant pas encore très connu, nous totalisons un nombre de huit recommandations dont seulement cinq ont été qualifiées par un seul de nos visiteurs. Notre agent artificiel se trouve alors dans la situation d'apprentissage suivante.

	« Cadre atypique » (C)	« Bicolore » (B)	« Vintage » (V)	« Like » (L)
Recommandation 1	1	0	0	0
Recommandation 2	1	0	1	1
Recommandation 3	1	1	0	0
Recommandation 4	1	1	1	?
Recommandation 5	0	0	0	0
Recommandation 6	0	0	1	?
Recommandation 7	0	1	0	?
Recommandation 8	0	1	1	0

Tableau 1 : Situation d'apprentissage 1

À partir du cas des recommandations 1, 2, 3, 5 et 8, nous souhaitons donc que notre agent artificiel prédise les préférences de notre visiteur pour les recommandations 4, 6 et 7. En d'autres termes, le problème d'apprentissage de notre agent artificiel consiste à trouver la fonction h^* lui permettant de classer les valeurs manquantes que représente chaque point d'interrogation.

Bien entendu, notre exemple d'apprentissage est extrêmement simple. Notre agent artificiel pourrait donc apprendre par cœur l'ensemble des fonctions permettant d'associer à chaque recommandation la préférence lui correspondant. Mais ce type d'apprentissage en extension ne nous convient pas puisqu'il ne pourrait pas être réalisé si nous ajoutons une dizaine de descripteurs binaires et si nous avons quelques milliers de visiteurs et quelques centaines de recommandations.

C'est donc ici que se trouve tout l'intérêt d'introduire l'espace des hypothèses \mathcal{H} . Car, c'est précisément à travers cet espace \mathcal{H} que notre agent artificiel va pouvoir décrire l'ensemble des partitions de \mathcal{X} dans un langage de description $\mathcal{L}_{\mathcal{H}}$. Ainsi, comme le symbolise le schéma 2, avec l'introduction de \mathcal{H} , notre agent artificiel devient capable de manipuler l'espace de description \mathcal{X} en intention.

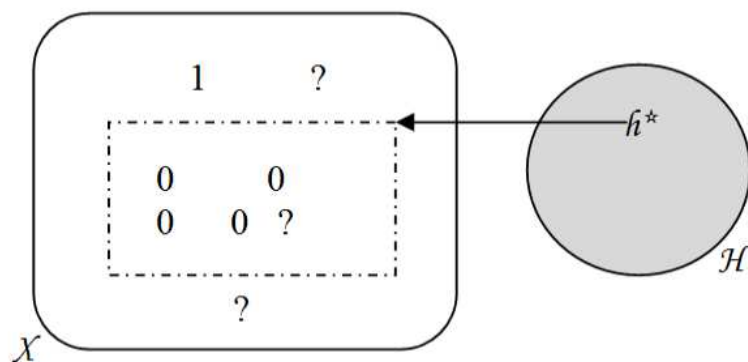


Schéma 2 : Situation d'apprentissage 1

Partant, comment notre agent artificiel peut-il faire pour prédire les préférences de notre visiteur sans faire de l'apprentissage par cœur ?

A priori, dans notre exemple, il ne peut rien faire. Pourquoi ? Tout simplement parce que, compte tenu des données entrantes χ , chaque hypothèse de \mathcal{H} renvoie à une conjonction bien spécifique de trois descripteurs binaires et que nous avons seulement cinq exemples dont les sorties y sont étiquetées. Par conséquent, notre agent artificiel ne peut pas réaliser d'induction logique.

Afin de permettre l'apprentissage de notre agent artificiel, nous sommes donc conduit à borner le langage de description $\mathcal{L}_{\mathcal{H}}$. Dans cette fiction, ayant une bonne connaissance du métier de vendeur de vélo de route et ayant découvert, au cours d'une analyse multifactorielle exploratoire réalisée à partir d'autres données, que l'attribut « Bicolore » (\mathcal{B}) ne pouvait pas être considéré comme un facteur significatif de préférence, nous décidons de ne pas considérer cet attribut. Notre agent se trouve alors dans la situation d'apprentissage suivante⁸.

	« Cadre atypique » (\mathcal{C})	« Vintage » (\mathcal{V})	« Like » (\mathcal{L})
Recommandation 1	1	0	0
Recommandation 2	1	1	1
Recommandation 3	1	0	0
Recommandation 5	0	0	0
Recommandation 8	0	1	0

Tableau 2 : Situation d'apprentissage 2

Dès lors, notre agent artificiel peut maintenant induire logiquement les préférences de notre visiteur :

- le vélo présenté sur la recommandation 4 est considéré comme beau ;
- le vélo présenté sur la recommandation 6 n'est pas considéré comme beau ;
- le vélo présenté sur la recommandation 7 n'est pas considéré comme beau⁹.

⁸ Ajoutons qu'en règle générale, dans un environnement big data, même si la quantité des descripteurs est très grande, il est possible de réaliser un apprentissage par induction compte tenu que le nombre de données entrantes χ est également très grand. Pour autant, il est toujours nécessaire de choisir un biais à travers ce que les spécialistes de l'apprentissage artificiel nomment le compromis biais-variance. Ce compromis peut être résumé ainsi : si le biais est trop grand, l'espace \mathcal{H} est trop restreint, la solution h^* risque alors d'être trop éloignée de la réalité ; mais, si la variance est trop grande, l'espace \mathcal{H} est trop large, la solution h^* risque de ne pas pouvoir être trouvée (Cornuéjols et Miclet, 2010). L'exemple que nous prenons est donc extrêmement simpliste et irréaliste, mais il permet de bien comprendre le problème du choix du biais d'apprentissage.

⁹ En d'autres termes, notre agent artificiel fait l'hypothèse que notre visiteur préfère les vélos qui ont un cadre atypique (\mathcal{C}) et qui sont vintage (\mathcal{V}).

	« Cadre atypique » (C)	« Bicolore » (B)	« Vintage » (V)	« Like » (L)
Recommandation 1	1	0	0	0
Recommandation 2	1	0	1	1
Recommandation 3	1	1	0	0
Recommandation 4	1	1	1	1
Recommandation 5	0	0	0	0
Recommandation 6	0	0	1	0
Recommandation 7	0	1	0	0
Recommandation 8	0	1	1	0

Tableau 3 : Situation d'apprentissage 3

Néanmoins, il est important de rappeler que la fonction cible h^* sous-tendue par le tableau 2 est biaisée puisqu'elle ne considère pas l'attribut « Bicolore » (B). Cet exemple permet ainsi de souligner la nécessité de bien connaître le jeu des données ou encore d'avoir une bonne connaissance métier lorsqu'on souhaite paramétrer l'apprentissage d'un agent artificiel¹⁰.

2.3 Le protocole d'apprentissage

Partant, après avoir établi le biais d'apprentissage, il faut définir le protocole d'apprentissage de l'agent artificiel de façon à cadrer les interactions qu'il entretient avec son environnement d'apprentissage et avec son oracle. Comme le montre le schéma 3, il existe deux grands types de protocole qui sont plus ou moins adaptés aux tâches de prédiction et aux tâches d'identification (Cornuéjols et Miclet, 2010).

¹⁰ Notons qu'il existe des techniques permettant d'effectuer de façon automatique la sélection des attributs à considérer dans le processus d'apprentissage artificiel. Auquel cas, la connaissance du jeu de données ou encore la connaissance métier est bien entendu beaucoup moins indispensable. Toutefois, bien que ces techniques soient très largement mobilisées par les marchands, il apparaît que ces derniers cherchent généralement à comprendre et maîtriser, au moins en partie, l'étape du choix du biais d'apprentissage.

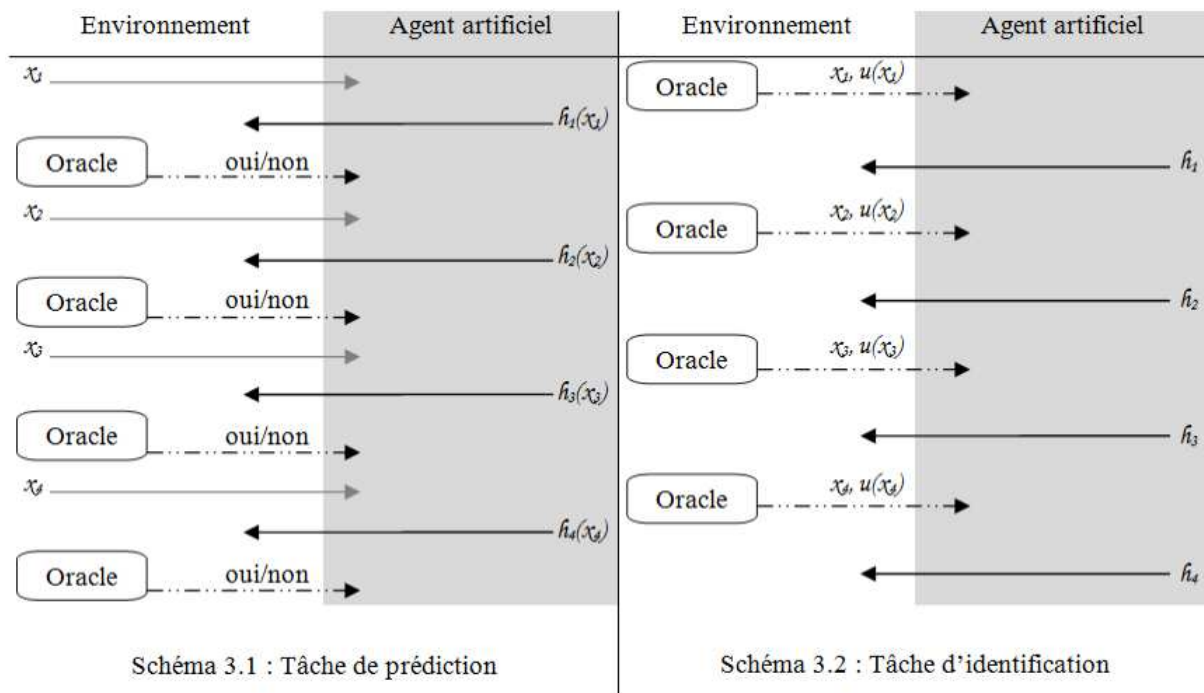


Schéma 3 : Protocoles d'apprentissage

Le schéma 3.1 illustre le protocole d'apprentissage par prédiction. À chaque fois que l'agent artificiel reçoit une description x , il propose une réponse y , et ensuite, l'oracle fournit la correction. Cette correction permet alors à l'apprenant d'améliorer progressivement la qualité de sa prédiction. Le schéma 3.2 illustre le protocole d'apprentissage par identification. Ici, à la différence du protocole d'apprentissage par prédiction, l'agent artificiel ne cherche pas à prédire la sortie y d'une entrée x mais à identifier la fonction f qui permet d'associer les données de l'ensemble $S = [(x_1, y_1), \dots, (x_n, y_n)]$. Bien entendu, une fois cette fonction identifiée à travers \hat{f}^* , l'agent artificiel peut réaliser un certain nombre de prédiction.

Ces deux protocoles d'apprentissage peuvent alors être réalisés en ligne ou hors ligne (Cornuéjols et Miclet, 2010). Dans le cas d'un protocole effectué hors ligne, l'ensemble des descriptions (x_1, x_2, \dots, x_n) sont fournies à l'agent artificiel en une seule fois. *A contrario*, dans le cas d'un protocole d'apprentissage en ligne, l'ensemble des descriptions (x_1, x_2, \dots, x_n) sont fournies à l'agent artificiel de façon séquentielle (généralement *au fil de l'eau*).

Dans notre cas fictif, il apparaît par exemple que le protocole d'apprentissage par prédiction et en ligne soit le plus adapté. En effet, en imaginant que notre site de vente de vélo de route ait fini par fonctionner et que nous ayons pu complexifier et affiner notre scénario d'apprentissage, par le biais des indices que sont les achats, les produits mis dans le panier et les produits les plus consultés, un tel protocole pourrait permettre à notre agent artificiel d'apprendre *au fil de l'eau* les préférences de nos visiteurs. Notre agent artificiel deviendrait ainsi capable de personnaliser l'environnement numérique de nos visiteurs en leur soumettant des recommandations susceptibles de les intéresser.

2.4 L'évaluation de l'apprentissage

Afin de finaliser le scénario d'apprentissage d'un agent artificiel, il est alors nécessaire d'établir un dispositif d'évaluation de l'apprentissage (Cornuéjols et Miclet, 2010).

Pour commencer, il faut élaborer un principe inductif permettant d'évaluer le risque empirique, c'est-à-dire, le risque d'erreur calculé à partir des données d'apprentissage. En référence à notre exemple, ce principe pourrait être élaboré en fonction des préférences

implicites de nos visiteurs que nous pourrions dégager à partir de leurs achats, de leur panier et des produits qu'ils consulteraient le plus fréquemment.

Ensuite, il faut appliquer ce même principe inductif à l'ensemble des données stockées hors-apprentissage afin d'évaluer le risque réel, c'est-à-dire, le risque d'erreur calculé à partir des données qui n'ont pas servi à l'apprentissage de l'agent artificiel¹¹.

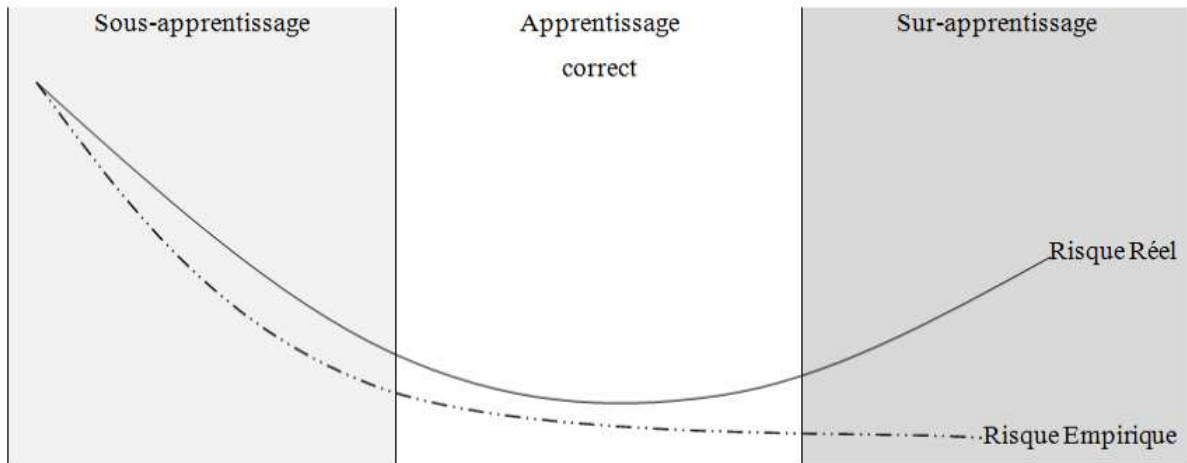


Schéma 4 : Évaluation de l'apprentissage

Partant, comme le montre le schéma ci-dessus, nous pourrions alors évaluer l'apprentissage de notre agent artificiel afin d'éviter les phénomènes de sous-apprentissage qui se manifestent par une trop grande généralisation et de sur-apprentissage qui se manifeste par une trop grande spécialisation.

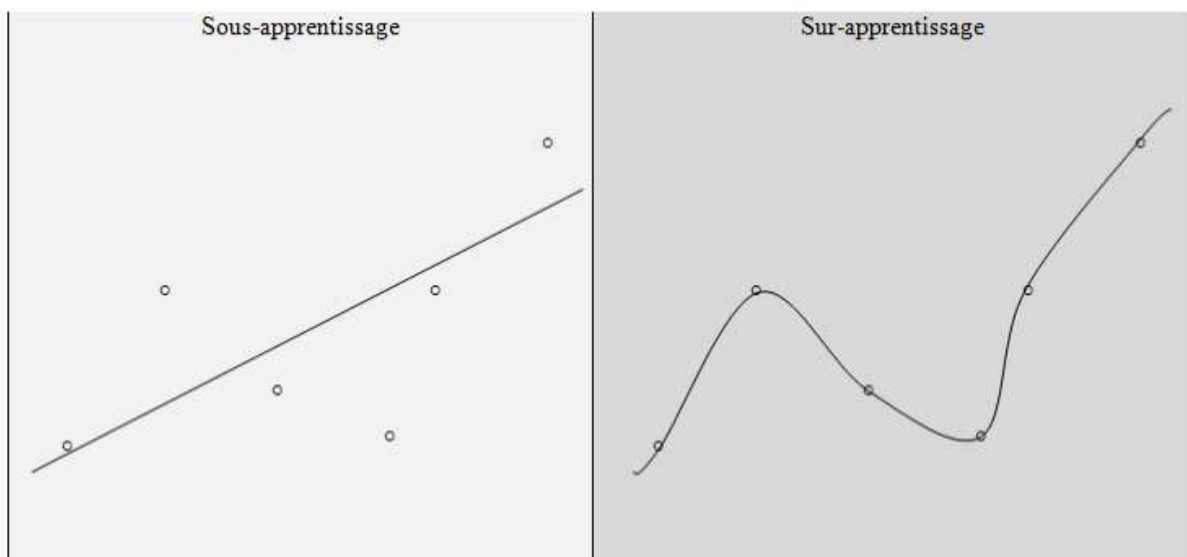


Schéma 5 : Sous-apprentissage et sur-apprentissage

¹¹ En effet, en apprentissage artificiel, il est généralement nécessaire de disposer de deux jeux de données : un jeu réservé à l'apprentissage et un autre permettant d'évaluer cet apprentissage (Cornuéjols et Miclet, 2010).

3 Les techniques d'apprentissage

L'apprentissage d'un agent artificiel fait donc l'objet d'un scénario plus ou moins détaillé qui implique un travail de paramétrage plus ou moins important.

Aussi, dans cette section, nous allons voir qu'il existe plusieurs techniques d'apprentissage qui, plus ou moins supervisées, permettent aux agents artificiels de développer des connaissances de façon plus ou moins autonomes.

Ici, notre questionnement est donc le suivant : quelles sont les grandes techniques d'apprentissage artificiel et, du point de vue des agents artificiels, quels sont les types de connaissances que ces techniques permettent de développer ?

Dans un premier temps, nous verrons que les techniques d'apprentissage par exploration favorisent avant tout le développement de connaissances symboliques dites supervisées (section 3.1). Ensuite, nous présentons les techniques d'apprentissage par optimisation qui encouragent cette fois-ci le développement de connaissances plutôt de type numérique et supervisé (section 3.2). Pour finir, nous montrons que les techniques d'apprentissage par découverte permettent de développer des connaissances symboliques et/ou numériques dites non-supervisées (section 3.3).

3.1 Les apprentissages par exploration

Il existe plusieurs formes d'apprentissages par exploration qui repose, le plus souvent, sur le modèle de l'espace des versions développé par Tom M. Mitchell en 1978 (Mitchell, 1997). Ici, nous nous concentrerons sur deux techniques d'apprentissage par exploration mobilisées par les marchands : la programmation logique inductive (*cf.* Ilic et *al.*, 2008) et l'inférence grammaticale (*cf.* Karampatziakis et *al.*, 2004 ; Korfiatis et Paliouras, 2008).

Comme nous l'avons vu dans la section 2, la programmation logique inductive¹² est une technique d'apprentissage symbolique supervisée qui peut permettre de développer un système de personnalisation basé sur le contenu. Par le biais de la programmation logique inductive, un agent artificiel peut en effet induire un ensemble de règles logiques à partir de différents exemples positifs et négatifs. En guise de deuxième illustration, imaginons cette fois-ci que les données issues de notre site de vente de vélo de route montrent en synthèse que :

- la clause « achat_pompe_vélo » (x) implique la clause « achat_rustine_vélo » (y) ;
- la clause « achat_vélo » (z) implique la clause « achat_rustine_vélo » (y) ;
- la clause « achat_vélo » (z) implique la clause « achat_pompe_vélo » (x).

Notre agent artificiel serait alors capable d'apprendre la règle suivante :

- pour tout achat d'un vélo (z) et de rustine à vélo (y) tel que l'achat d'un vélo (z) implique celui de rustine à vélo (y), il existe un achat de pompe à vélo (x) tel que l'achat de pompe à vélo (x) implique celui de rustine à vélo (y) et tel que l'achat d'un vélo (z) implique celui d'une pompe à vélo (x).

Notons que cette règle pourrait ainsi permettre à notre agent artificiel de recommander l'achat d'une pompe à vélo à un visiteur qui viendrait, par exemple, d'acheter un vélo et des rustines à vélo.

¹² Notons que c'est Gordon D. Plotkin (1970) qui a posé les bases de la programmation logique inductive.

L'inférence grammaticale¹³ est également une technique d'apprentissage symbolique supervisée. Elle est utilisée par les marchands afin de dégager des groupes d'actions (i.e. : *patterns*) caractéristiques des parcours informationnels de leurs visiteurs. En effet, l'inférence grammaticale permet d'induire des motifs correspondant à des stratégies d'action. Ainsi, notre agent artificiel pourrait apprendre les stratégies de découvertes (SV_D) et de recherches d'information (SV_R) de nos visiteurs afin de développer, en fonction d'un critère d'optimisation préalablement établie (e.g. : un indice de désorientation à minimiser), la stratégie de recommandation (SAA_C et SAA_S) présentée dans le schéma 6.

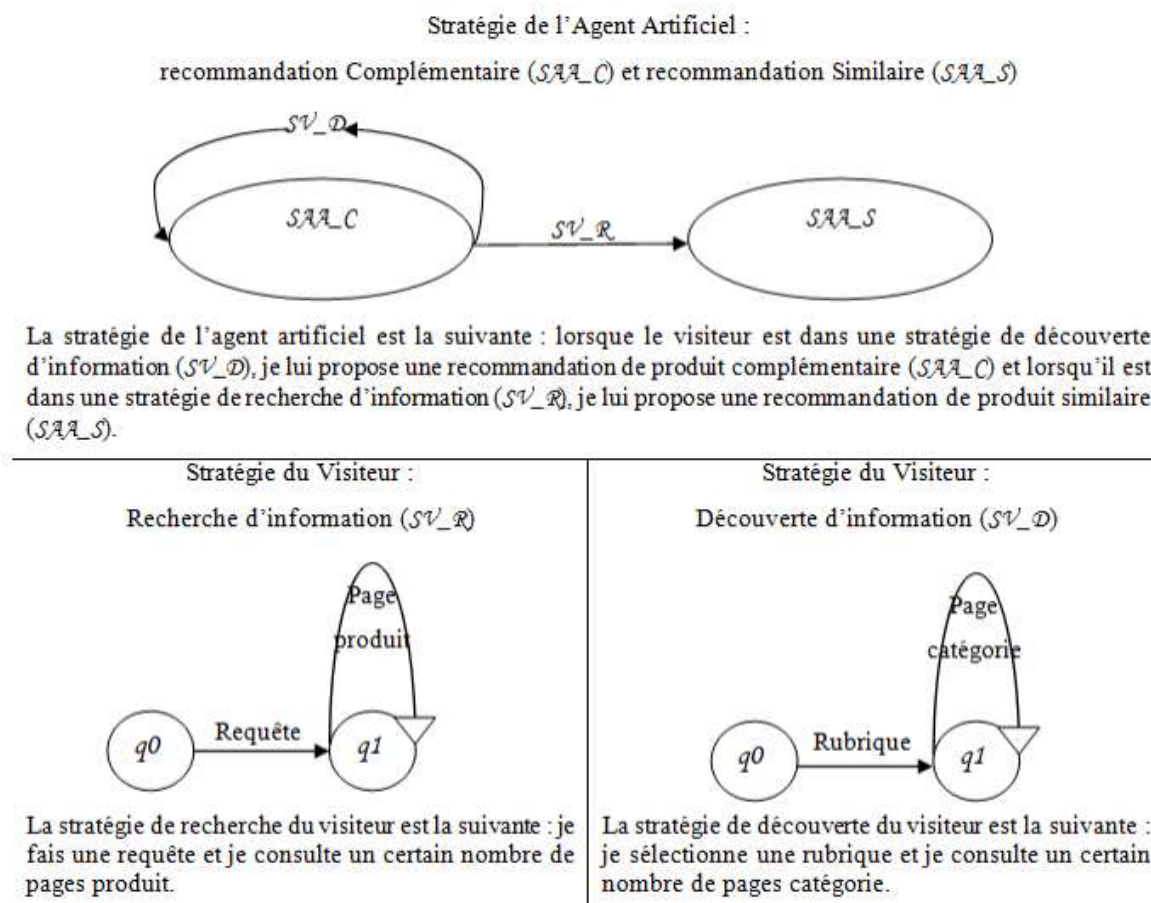


Schéma 6 : Apprentissage par inférence grammaticale

3.2 Les apprentissages par optimisation

Comme dans les cas des techniques d'apprentissage par exploration, les techniques d'apprentissage par optimisation sont nombreuses. Cependant, ces techniques d'apprentissage ne sont pas dites symboliques mais numériques. Ici, nous proposons de présenter deux techniques relativement classiques dans le domaine de l'apprentissage artificiel : la régression linéaire et le réseau de neurones. Ajoutons que ces deux techniques sont souvent mobilisées dans le domaine du e-commerce pour réaliser différents systèmes de filtrage collaboratif et/ou basé sur le contenu (cf. Vucetic, 2005 ; Vassiliou et al., 2006).

¹³ Pour Colin de la Higuera (2005), les auteurs ayant contribué à faire émerger l'inférence grammaticale au sein de l'intelligence artificielle sont nombreux. Néanmoins, il souligne que les articles de King-Sun Fu (1974) et de King-Sun Fu et Taylor L. Booth (1975) sont les premiers à proposer une définition approfondie de ce domaine de recherche.

La régression linéaire¹⁴ est la technique d'apprentissage par optimisation la plus connue. Et, comme l'illustre le schéma 7, à partir de cette technique, notre agent artificiel pourrait par exemple apprendre que plus nos visiteurs ont tendance à interagir avec leur panier (O_p) et plus ils ont tendance à effectuer des parcours informationnelles redondants (\mathcal{R}_i).



O_p : fréquence d'Ouverture de la page panier.

\mathcal{R}_i : Indice de Redondance informationnelle calculant le ratio du nombre de page identique ouverte (pi) sur le nombre de pages totales ouvertes (pt), soit : $\mathcal{R}_i = pi/pt$.

Schéma 7 : Apprentissage par régression linéaire

En référence à l'exemple précédent et en admettant qu'un indice de redondance informationnelle (\mathcal{R}_i) élevé couplé à une forte fréquence d'ouverture de la page panier (O_p) indique une posture de recherche d'information, notre agent artificiel pourrait alors automatiser la décision d'activation du système de recommandation de produits similaires à chaque fois qu'un visiteur ouvre plus d'une fois la page panier. En outre, si la régression linéaire permet aux agents artificiels d'apprendre à réaliser facilement et rapidement des prédictions efficaces à partir de variables quantitatives, elle constitue néanmoins une technique d'apprentissage qui n'est pas très expressive.

Le réseau de neurones¹⁵ est, en quelque sorte, la symétrie inversée de la régression linéaire. Il est composé d'un plus ou moins grand nombre d'unités élémentaires (i.e. : les perceptrons ; Rosenblatt, 1962) capables d'apprendre des fonctions f de décisions simples. Et, le principe d'apprentissage d'un tel réseau est le suivant :

- plus deux perceptrons s'activent l'un et l'autre, plus ils sont liés l'un à l'autre (Hebb, 1949).

¹⁴ La théorie de la discrimination linéaire est très ancienne puisqu'elle remonte aux premières procédures de classification développées par Ronald A. Fisher (1932). Cependant, elle attira l'attention des spécialistes de l'intelligence artificielle quelques années plus tard avec les travaux de Franck Rosenblatt sur le perceptron (1962 ; cf. l'apprentissage par réseau de neurones).

¹⁵ Les travaux réalisés par Warren S. McCulloch et Walter H. Pitts (1943) sont souvent cités comme au fondement des réseaux de neurones.

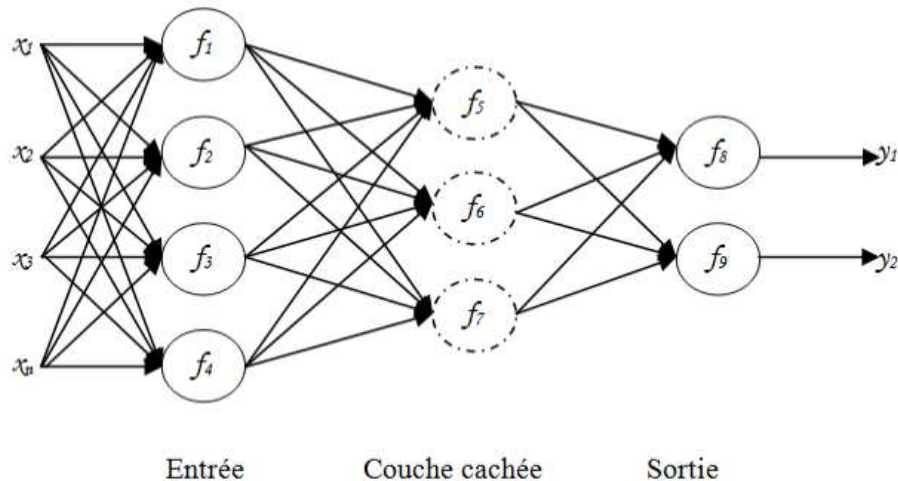


Schéma 8 : Apprentissage par réseau de neurones

Compte tenu de la multiplicité des schèmes d'associations permis par les réseaux de neurones, ces derniers ont alors de très bonnes capacités d'apprentissages dans les domaines de la généralisation et de la catégorisation. Cependant, si les réseaux de neurones permettent de rendre compte d'une certaine complexité, leur apprentissage, qui demande une grande quantité de données, est relativement long et difficile.

Pour autant, à partir d'une série de vingt variables ($v_1, v_2, v_3, \dots, v_{20}$; cf. les fonctions d'entrée) caractérisant les parcours informationnels et les comportements d'achat de plusieurs milliers d'anciens clients (cf. les exemples χ), la technique du réseau de neurone pourrait permettre à notre agent artificiel d'apprendre à reconnaître le profil de chacun de nos nouveaux clients à l'aide des sortie y que sont les étiquettes « Explorateurs Fidèles (EF) », « Explorateurs Infidèles (EI) », « Chercheurs Fidèles (CF) » et « Chercheurs Infidèles (CI) ». Il pourrait ainsi personnaliser l'environnement numérique de nos nouveaux visiteurs en décidant d'activer ou non le système de recommandation de produits similaires (SAA_S) ou complémentaires (SAA_C) tout en ajoutant un paramètre de fidélisation (e.g. : une promotion) aux clients les moins fidèles.

3.3 Les apprentissages par découverte

Les techniques d'apprentissage par découverte sont aujourd'hui très répandues dans le domaine du commerce numérique (cf. Amatriain et al., 2011 ; Ben Schafer, 2009). D'une manière générale, ce type d'apprentissage renvoie au problème de la classification non-supervisée. Il existe plusieurs méthodes de classification non-supervisée : par exemple, celles fondés sur les distances et celles fondées sur les hiérarchies. Ici, nous présentons seulement un exemple de classification non-supervisée fondée sur les distances.

Comme toutes les autres formes de classification non-supervisée¹⁶, celles fondées sur les distances permettent, en quelque sorte, de réaliser le travail d'étiquetage effectué par l'oracle. Par exemple, par le biais de l'algorithme des k-moyennes ou encore des cartes auto-organisatrices de Kohonen, un agent artificiel peut dégager la division « naturelle » des données χ en classe y . Plus concrètement, en référence à notre exemple précédent, notre

¹⁶ Les techniques de classification trouvent bien entendu leurs origines dans les travaux de Karl Pearson (1894). Cependant, il ressort que le concept de fouille de données a été défini et diffusé part l'ouvrage de Gregory Piatetsky-Shapiro et William J. Frawley (1991).

apprenant artificiel pourrait apprendre à classifier en toute autonomie l'ensemble de nos visiteurs afin de dégager les quatre groupes que sont les explorateurs fidèles (EF), les explorateurs infidèles (EI), les chercheurs fidèles (CF) et les chercheurs infidèles (CI).

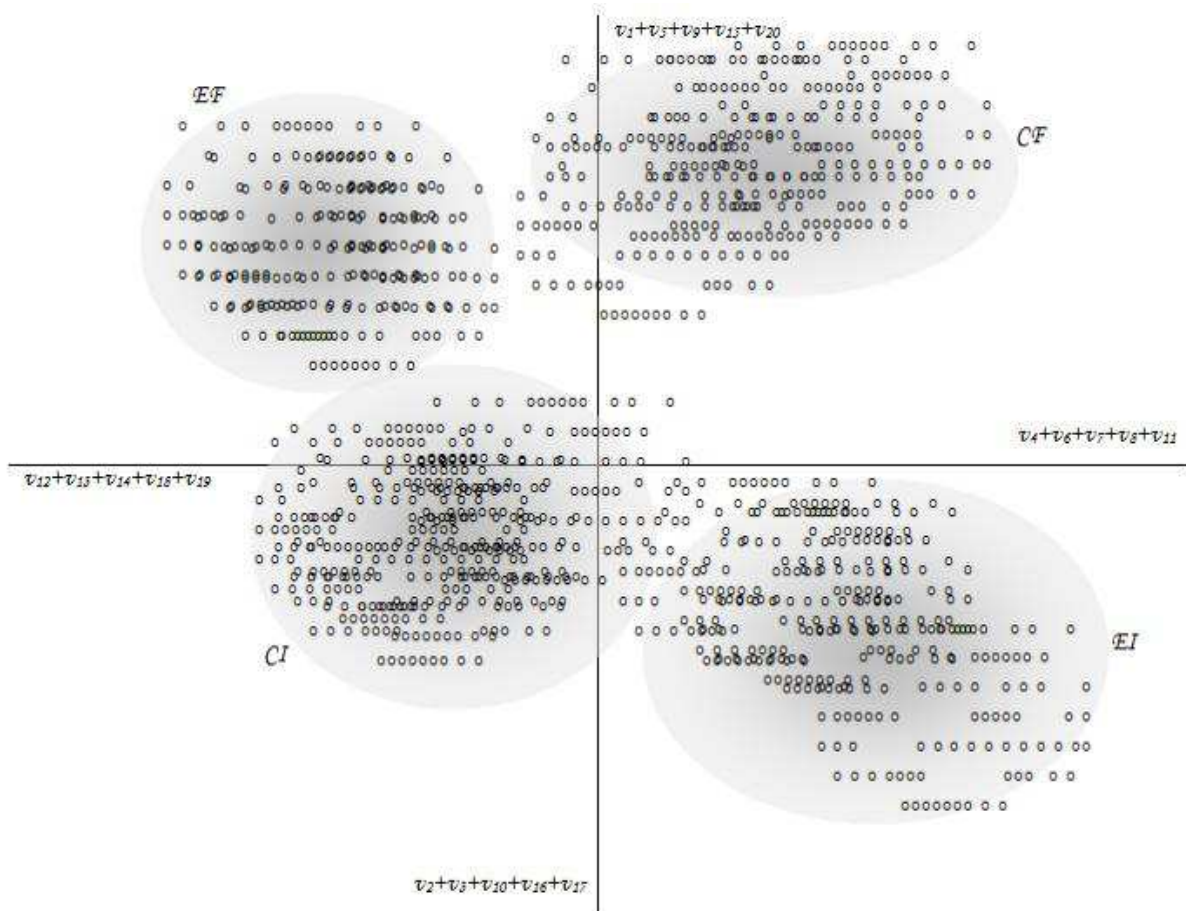


Schéma 9 : Apprentissage par classification non-supervisée

Ainsi, notre agent artificiel pourrait définir par lui-même les quatre étiquettes « Explorateurs Fidèles (EF) », « Explorateurs Infidèles (EI) », « Chercheurs Fidèles (CF) » et « Chercheurs Infidèles (CI) » de façon à apprendre le réseau de neurones présenté précédemment et ainsi pouvoir reconnaître et prédire les profils de chacun de nos nouveaux clients.

CONCLUSION

Avec le développement toujours croissant des TIC, nous sommes tous conduit à produire une grande quantité de trace d'usages (i.e. : les big data) que les marchands stockent et traitent afin de personnaliser automatiquement nos environnements numériques marchands (cf. section 1). Pour ce faire, ces derniers effectuent deux grands types d'activité :

- la première, qui n'est pas toujours très soignée, consiste à vérifier la qualité épistémologique des big data ;
- la seconde, qui fait l'objet d'un travail parfois considérable, consiste à développer des agents artificiels capables d'inférer des connaissances à partir de ces big data.

Afin de permettre l'apprentissage des agents artificiels, les marchands sont amenés à développer un problème et un scénario d'apprentissage (cf. section 2), puis, à choisir (voire parfois à combiner) la(les) technique(s) d'apprentissage(s) qui est(ont) la(les) plus adaptée(s) à ce dernier (cf. section 3).

En ce sens, afin d'élaborer, par exemple, un système de recommandation, les marchands doivent réaliser un travail de conception qui s'apparente à ce que Pierre-Michel Menger (2014) nomme le travail créateur¹⁷. Et, comme nous nous sommes attaché à le montrer dans la section 2 et 3, c'est très précisément à travers ce travail que les marchands conçoivent des agents artificiels qui constituent, en quelque sorte, des symbiotes intégrés des cognitions des consommateurs (Brangier et *al.*, 2010). Ces symbiotes instaurent ainsi une certaine forme de coévolution entre l'utilisateur et la machine, puisque, comme l'avait déjà anticipé Joseph C.R. Licklider en 1960, les apprenants artificiels instituent finalement une dépendance cognitive et dynamique entre ces deux acteurs/actants. Notons alors que, du point de vue des marchands, la maîtrise de la qualité de cette relation est lourde de conséquences puisqu'elle doit leur permettre de personnaliser les interactions qu'ils entretiennent avec une masse de consommateurs parfois très hétérogènes.

Pour autant, cette relation symbiotique n'est pas symétrique. Car, si elle permet effectivement de mieux considérer les capacités de créativité des usagers (Certeau, 1990), elle le fait selon une logique d'exploitation. En effet, à travers les big data et les techniques d'apprentissage artificiel, il s'agit bel et bien, pour les marchands, d'exploiter les logiques d'exploration développées par les consommateurs durant leurs activités numériques (March, 1991 ; Mallard, 2011). Par conséquent, si, du point de vue de l'offre, cette relation symbiotique est intéressante compte tenu qu'elle favorise un modèle de gestion plus ambidextre (Mothe et Brion, 2008), elle n'en reste pas moins une stratégie d'exploitation.

En conclusion, nous coproduisons nos environnements numérique marchand à travers :

- la créativité dont nous faisons preuve durant les usages que nous faisons quotidiennement des TIC ;
- la créativité dont font preuve les marchands pour développer les agents artificiels qui vont apprendre nos appétences, nos centres d'intérêts et nos pratiques de consommation à partir des usages créatifs que nous faisons des TIC.

Néanmoins, compte tenu que les big data sont des traces d'usages qui ne sont généralement pas sollicitées et que nous n'avons aucune possibilité de connaître et encore moins de participer au paramétrage des apprenants artificiels, il semble que cette relation symbiotique que nous entretenons aujourd'hui avec notre environnement numérique marchand ne puisse pas, à notre sens, être considérée comme une cocréation. En effet, il ne faut pas oublier qu'à travers les différentes technologies de personnalisation, la coproduction des environnements numériques marchands est d'abord une création des marchands. Et surtout, il ne faut pas négliger le fait que les agents artificiels sont créés (et donc se développent) avant tout en fonction des représentations, des incertitudes et des intérêts des seuls marchands.

¹⁷ cf. par exemple l'entretien que *Libération* a réalisé avec David Bessis en janvier 2012 : « L'intuition mathématique, c'est comme une transe ». Rappelons que David Bessis est le fondateur de *tyniclues*, une société qui propose des services de personnalisation des environnements numériques marchands. Et, dans cet article, David Bessis exprime bien quelques uns des aspects créatifs que recouvre la pratique des mathématiques. En ce sens, David Bessis rejoint les autres témoignages que nous avons pu recueillir dans d'autres travaux et qui montre que la conception de technologies algorithmiques relève souvent d'un travail de création complexe et tâtonnant qui, mêlant intuition et une certaine recherche de « pureté », demande une forte motivation intrinsèque (cf. Menger, 2009).

BIBLIOGRAPHIE

Akrich M.

1989, *La construction d'un système socio-technique. Esquisse pour une anthropologie des techniques*, in *Anthropologie et Sociétés*, n° 2, vol. 13, pages 31-54.

Amatriain X., Jaimes A., Oliver N., Pujol J.M.

2011, *Data Mining Methods for Recommender Systems*, in *Recommender Systems Handbook*, pages 39-71.

Baiocchi M.C., Forest D.

2014, *L'usager comme autorité cognitive. Perspectives théoriques sur les systèmes de recommandation*, in *Les Cahiers du numérique*, vol. 10, pages 127-157.

Ben Schafer J.

2009, *The Application of Data-Mining to Recommender Systems*, in *Encyclopedia of Data Warehousing and Mining*, pages 45-50.

Boullier D.

2004, *L'habitèle comme extension du domaine de la personne*, in *Consommation et société*, n°4, pages 1-12.

Brangier E., Dufresne A., Hammes-Adelé S.

2009, *Approche symbiotique de la relation humain-technologie : perspectives pour l'ergonomie informatique*, in *Le travail humain*, n° 4, vol. 72, pages 333-353.

Certeau M. (de)

1990, *L'invention du quotidien, tome 1 : Arts de faire*, Gallimard.

Cochoy F.

2004, *Introduction. La captation des publics entre dispositifs et dispositions, ou le petit chaperon rouge revisité*, in Cochoy F., *La captation des publics*, PUM.

Cornuéjols A., Miclet L.

2010, *Apprentissage artificiel : Concepts et algorithmes*, Eyrolles.

Desrosières A.

2008, *L'argument statistique. Pour une sociologie historique de la quantification*, PEM.

Fisher R.A.

1932, *Statistical Methods for Research Workers*, Oliver and Boyd.

Fu K.-S.

1974, *Syntactic Methods in Pattern Recognition*, Academic Press.

Fu K.-S., Booth T.L.

1975, *Grammatical Inference: Introduction and survey*, in *Systems, Man, and Cybernetics*, *IEEE Transaction*, n° 4, vol. 5, pages 409-423.

Gitelman L.

2006, *« Raw Data » Is an Oxymoron*, MIT Press.

Higuera C. (de la)

2005, *A bibliographical study of grammatical inference*, in *Pattern Recognition*, n° 9, vol. 38, pages 1332-1348.

Hebb D.O.

1949, *The Organization of Behavior: A Neuropsychological Theory*, Wiley.

Ilic M., Leite J., Slota M.

2008, *Explicit Dynamic User Profiles for a Collaborative Filtering Recommender System*, in *Advances in Artificial Intelligence*, vol. 5290, pages 352-361.

Jeanneret Y.

2013, *Faire trace : un dispositif de présentation du social*, in *Intellectica*, n° 59, vol. 1, pages 41-64.

2011, *Complexité de la notion de trace : de la traque au tracé*, in Galinon-Méléneq B., *L'homme-trace : perspectives anthropologiques des traces contemporaines*, CNRS.

Kessous E.

2012, *L'attention au monde. Sociologie des données personnelles à l'ère numérique*, Armand Colin.

Korfiatis G., Paliouras G.

2008, *Modeling Web Navigation Using Grammatical Inference*, in *Applied Artificial Intelligence*, n° 22, vol. 1, pages 116-138.

Karampatziakis N., Paliouras G., Pierrakos D., Stamatopoulos P.

2004, *Navigation Pattern Discovery Using Grammatical Inference*, in *Grammatical Inference: Algorithmes and applications*, vol. 3264, pages 187-198.

Latour B.

1993, *Le « pédofil » de Boa Vista ou la référence scientifique*, in Latour B., *La clef de Berlin et autres leçons d'un amateur de science*, La Découverte.

Leleu-Merviel S.

2013, *Traces, information et construits de sens. Déploiement de la trace visuelle de la rétention indicielle à l'écriture*, in *Intellectica*, n° 59, vol. 1, pages 65-88.

Licklider J.C.R

1960, *Man-Computer Symbiosis*, in *Transactions on Human Factors in Electronics*, vol. HFE-1, pages 4-11.

Mallard, A.

2011, *Explorer les usages : un enjeu renouvelé pour l'innovation des TIC*, in Denouel J., Granjou F., *Communiquer à l'ère numérique. Regards croisés sur la sociologie des usages*, Presses des Mines.

March J.G.

1991, *Exploration and Exploitation in Organizational Learning*, in *Organization Science*, n° 1, vol. 2, pages 71-87.

Marcus G.E.

1995, *Ethnography in/Of the World System: The Emergence of Multi-sited Ethnography*, in *Annual Review of Anthropology*, vol. 24, pages 95-117.

McCulloch W.S., Pitt W.H.

1943, *A logical calculus of the ideas immanent in nervous activity*, in *Bulletin of Mathematical Biophysics*, vol. 5, pages 115-133.

Ménard M.

2014, *Systèmes de recommandation de biens culturels. Vers une production de conformité ?*, in *Les Cahiers du numérique*, vol. 10, pages 69-94.

Menger, P.-M.

2014, *The Economy of Creativity. Art and Achievement Under Uncertainty*, Harvard University Press.

Mille A.

2013, *De la trace à la connaissance à l'ère du Web. Introduction au dossier*, in *Intellectica*, n° 59, vol. 1, pages 7-28.

Mitchell T.M.

1997, *Machine Learning*, McGraw Hill.

Mothe C., Brion S.

2008, *Innovation : exploiter ou explorer ?*, in *Revue Française de Gestion*, n° 187, pages 101-108.

Pearson K.

1984, *Contributions to the Mathematical Theory of Evolution*, in *Philosophical Transactions of the Royal Society of London*, pages 71-110.

Piatetsky-Shapiro G., Frawley W.J.

1991, *Knowledge Discovery in Databases*, AAAI/MIT Press.

Plotkin G.D.

1970, *A note on induction generalization*, in *Machine Intelligence*, vol. 5, pages 165-179.

Rosenblatt F.

1962, *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*, Spartan.

Vassiliou C., Stamoulis D., Martakos D., Athanassopoulos S.

2006, *A recommender system framework combining neural networks and collaborative filtering*, in *IMCAS'06 Proceedings of the 5th WSEAS international conference on Instrumentation, measurement, circuits and systems*, pages 285-290.

Vayre J.S.

2013, *Le big data et la relation client. Quand les traces numériques organisent l'échange marchand*, in *12^e Journées Normandes de Recherche sur la Consommation : Société et consommation*, pages 1-19.

2014, *Manipuler les données. Documenter le marché. Les implications organisationnelles du mouvement big data*, in *Les Cahiers du Numérique*, n° 1, vol. 10, pages 95-125.

Vucetic S., Obradovic Z.

2005, *Collaborative Filtering Using a Regression-Based Approach*, in *Knowledge and Information Systems*, n° 1, vol. 7, pages 1-22.