



HAL
open science

Classification de patterns de catégorisation chez l'humain par deux modèles d'apprentissage

Giulia Mezzadri, Patricia Reynaud-Bouret, Thomas Laloë, Fabien Mathy

► **To cite this version:**

Giulia Mezzadri, Patricia Reynaud-Bouret, Thomas Laloë, Fabien Mathy. Classification de patterns de catégorisation chez l'humain par deux modèles d'apprentissage. JDS 2020: 52-ièmes journées de Statistique de la Société Française de Statistique (SFdS), Jun 2021, Nice (en ligne), France. hal-03508398

HAL Id: hal-03508398

<https://hal.science/hal-03508398>

Submitted on 3 Jan 2022

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.

CLASSIFICATION DE PATTERNS DE CATÉGORISATION CHEZ L'HUMAIN PAR DEUX MODÈLES D'APPRENTISSAGE

Giulia Mezzadri ¹ & Patricia Reynaud-Bouret ¹ & Thomas Laloë ¹ & Fabien Mathy ²

¹ *Laboratoire J.A. Dieudonné, CNRS UMR 7351, Nice*

giulia.mezzadri@univ-cotedazur.fr

patricia.reynaud-bouret@univ-cotedazur.fr

thomas.laloe@univ-cotedazur.fr

² *Bases, Corpus, Langage, CNRS UMR 7320, Nice*

fabien.mathy@univ-cotedazur.fr

Résumé. La catégorisation est un processus cognitif qui permet de regrouper des objets ayant une ou plusieurs caractéristiques en commun. Un des problèmes majeurs dans ce domaine est la multitude des modèles cognitifs qui ont comme but de reproduire les performances de participants humains dans une tâche de catégorisation. Nous nous proposons comme objectif de comparer deux modèles influents en psychologie : ALCOVE et Component-Cue, tous deux fondés sur une structure de réseaux de neurones artificiels. Après une présentation de ces deux modèles d'apprentissage, nous détaillerons la méthode que nous avons utilisée pour déterminer lequel de ces modèles s'ajuste au mieux à un jeu de données expérimentales. Nous montrerons enfin que la moitié des participants suit une logique plus proche d'ALCOVE, tandis que l'autre moitié suit une logique plus proche de Component-Cue.

Mots-clés. Cognition, Neurosciences, apprentissage de catégories, ALCOVE, réseau de neurones artificiels, Component-Cue, estimation par maximum de vraisemblance, méthode des moindres carrée.

Abstract. Categorization is a cognitive process that allows the grouping of objects with one or more common features. One of the major problems in this field is the multitude of cognitive models which aim to reproduce the performances of human participants in a categorization task. We aim to compare two influential models in psychology : ALCOVE and Component-Cue, both based on a structure of artificial neural networks. After a presentation of these two learning models, we will present the method we used to determine which of these models best fits an experimental dataset. We will finally show that half of the participants follow a logic closer to ALCOVE, while the other half follows a logic closer to Component-Cue.

Keywords. Cognition, neuroscience, categorization learning, ALCOVE, Component-Cue, artificial neural network, maximum likelihood estimation (MLE), Sum of squared differences (SSD).

1 Introduction

En observant le ciel, nous étiquetons immédiatement le temps comme étant beau ou mauvais ; une musique peut-être classée comme un morceau de jazz, folk, rap, ou autres styles ; en lisant un texte nous l’identifions comme un article de journal, un roman, une poésie ou autre encore. La catégorisation s’intéresse à l’ensemble de ce type de processus cognitifs qui permettent à un apprenant, qu’il soit un humain, un animal ou un algorithme, de placer des objets similaires dans un même groupe.

Un des problèmes majeurs dans ce domaine est la multitude et la variété des modèles qui ont comme but de reproduire les performances de participants humains dans une tâche de catégorisation. Un objectif est donc celui de comparer les modèles de catégorisation les plus influents en psychologie cognitive. Avant de se plonger dans cette analyse, il est important de comprendre en quoi consiste une expérience de catégorisation.

Dans une expérience de catégorisation, l’expérimentateur choisi un ensemble d’objets qu’il place arbitrairement dans deux catégories. L’objectif du participant est de comprendre et d’assimiler la règle de catégorisation choisie à travers un apprentissage par essai-erreur. La tâche de catégorisation est divisée en deux phases : une première phase, appelée phase d’apprentissage, dans laquelle le participant est censé apprendre la règle de catégorisation choisie par l’expérimentateur ; et une deuxième phase, appelée phase de transfert, dans laquelle l’introduction de nouveaux objets permet de tester la généralisabilité de l’apprentissage du participant.

Étant donné la différence de nature des phases d’apprentissage et de transfert, il y a des modèles, dits de transfert, qui sont plus adaptés à la phase de transfert, et d’autres, dits d’apprentissage, qui sont plus adaptés à la phase d’apprentissage. Les modèles de transfert sont utiles pour reproduire les performances finales des participants (celles obtenues à la fin du processus d’apprentissage), tandis que les modèles d’apprentissage peuvent rendre compte à la fois de la dynamique d’apprentissage et de la phase de transfert. Ici, nous nous focalisons sur deux modèles d’apprentissage qui ont montré dans le passé leur qualité d’ajustement aux données humaines (voir [3] et [4]) : ALCOVE et Component-Cue. Ces modèles sont tous deux fondés sur des réseaux de neurones et utilisent des conditions de similarité.

Après une présentation des spécificités des deux modèles d’apprentissage considérés, nous présenterons la méthode utilisée pour sélectionner le modèle qui s’adapte au mieux aux données. Nous montrerons enfin les résultats obtenus sur données simulées et expérimentales.

2 ALCOVE et Component-Cue

ALCOVE et Component-Cue sont deux modèles d’apprentissage basés sur une structure de réseau de neurones artificiels (Figure 1). Cette structure, conceptualisée pour reproduire le fonctionnement de l’esprit humain dans une tâche de catégorisation, est

constituée de trois parties : un premier nœud qui reçoit le stimulus (input), des nœuds centraux qui élaborent l’information, et une dernière couche de nœuds (outputs) qui est reliée aux précédents à travers des poids (spécifiques à chaque catégorie) et qui produit une sortie pour chacune des catégories possibles. À chaque fois que le réseau reçoit un stimulus, les nœuds s’allument selon une logique propre au modèle et la somme pondérée de tous les nœuds produit les sorties des catégories considérées. Les poids qui relient les

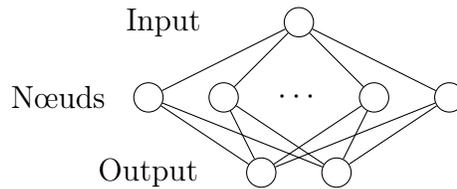


FIGURE 1 – Structure de réseau de neurones commune à ALCOVE et Component-Cue.

nœuds centraux à ceux de sortie sont mis à jour après chaque réception d’un stimulus. Pour mettre à jour les poids, les modèles utilisent un algorithme de descente de gradient qui cherche à minimiser l’écart entre la réponse du participant et les sorties du réseau. En particulier, cette étape rend le processus d’apprentissage dépendant de tout le passé, en augmentant ainsi la difficulté à analyser statistiquement ce type de modèles.

Les nœuds centraux d’ALCOVE et de Component-Cue élaborent l’information différemment. Dans ALCOVE chaque nœud code un objet de l’expérience et son activation dépend de la similarité entre l’objet associé au nœud et le stimulus d’input. Au contraire, dans Component-Cue chaque nœud code une des caractéristiques possibles de l’objet et son activation se base sur la présence ou l’absence de la caractéristique considérée dans le stimulus d’input. Si on dénote A et B les deux catégories dans lesquelles nous pouvons classifier les objets, alors pour les deux modèles la sortie associée à la catégorie A au temps t pour l’objet x est la suivante :

$$O_A^{(t)}(x) = \sum_{i=1}^n a_i(x) \cdot w_{iA}^{(t)},$$

où n représente le nombre de nœuds, $a_i(x)$ dénote l’activation du i -ème nœud suite à la réception de l’input x , et $w_{iA}^{(t)}$ représente le poids au temps t qui relie le i -ème nœud au nœud de sortie de la catégorie A . Au contraire, la probabilité de classifier un objet dans une des deux catégories revêt une forme différente dans les deux modèles. Pour ALCOVE, la probabilité de catégoriser l’objet x dans la catégorie A est donnée par :

$$\mathbb{P}(A | x, \mathcal{H}_{t-1}) = \frac{O_A^{(t)}(x) + b}{O_A^{(t)}(x) + O_B^{(t)}(x) + 2b},$$

tandis que pour Component-Cue, elle est donnée par :

$$\mathbb{P}(A | x, \mathcal{H}_{t-1}) = \frac{e^{\phi \cdot O_A^{(t)}(x)}}{e^{\phi \cdot O_A^{(t)}(x)} + e^{\phi \cdot O_B^{(t)}(x)}},$$

où b et ϕ sont deux paramètres fixes à estimer, et \mathcal{H}_{t-1} représente le passé du processus d'apprentissage jusqu'au temps $t - 1$.

Il est important de souligner que ALCOVE et Component-Cue peuvent aussi être utilisés sur des données provenant d'une phase de transfert. Pour ce faire, il suffit de fixer les poids du réseau de neurones. Pour citer un exemple, dans l'article [3], ALCOVE et Component-Cue ont été utilisés dans leur forme dynamique de réseau de neurones sur des données d'apprentissage, et ensuite dans leur forme statique sur des données de transfert.

3 Comment compare-t-on deux modèles d'apprentissage ?

Pour déterminer le modèle qui s'ajuste au mieux aux données, nous exploitons la division en deux phases, d'apprentissage et de transfert, caractéristique des expériences de catégorisation. La phase d'apprentissage est utilisée pour l'estimation des paramètres du modèle considéré, tandis que la phase de transfert est utilisée pour calculer l'erreur entre les réponses du participant et les prévisions du modèle. La phase d'apprentissage permet d'estimer les paramètres, tandis que la phase de transfert constitue notre test. L'estimation des paramètres du modèle est faite par maximum de vraisemblance (maximum likelihood estimation, MLE) :

$$\hat{\theta} \in \arg \min_{\theta} \mathcal{L}(\theta; \mathcal{D}_A),$$

où \mathcal{L} représente la vraisemblance du modèle considéré et \mathcal{D}_A représente les données en phase d'apprentissage. Au contraire, le calcul de l'erreur entre les prévisions et les réponses des participants est fait en utilisant la méthode des moindres carrés (sum of squared differences, SSD) :

$$\text{SSD} = \sum_{i \in \mathcal{D}_T} \left(p_i^{\hat{\theta}} - r_i \right)^2,$$

où, $p_i^{\hat{\theta}}$ dénote la probabilité donnée par le modèle de classifier le stimulus i dans la première catégorie (cette probabilité est obtenue en utilisant le vecteur des paramètres estimés $\hat{\theta}$), r_i dénote la réponse du participant au temps i et \mathcal{D}_T dénote les données en phase de transfert. L'utilisation du calcul de la SSD comme critère pour déterminer le modèle qui s'ajuste au mieux aux données est motivé par son utilisation en psychologie (voir [3] et [4]).

4 Application aux données simulées

Avant d'appliquer la méthode décrite ci-dessus aux données expérimentales, nous devons nous assurer que, quand nous l'appliquons aux données simulées, nous sommes bien capables de retrouver le modèle avec lequel nous avons créé les données. Pour ce faire, nous simulons des données artificielles avec un modèle donné et nous appliquons la méthode illustrée dans le paragraphe précédent afin d'établir, parmi tous les modèles, celui qui s'ajuste au mieux aux données simulées. Si nous retrouvons bien le modèle avec lequel nous avons simulées les données artificielles, nous pouvons affirmer que la méthode permet de discriminer les modèles.

Pour simuler les données artificielles nous avons utilisé les paramètres trouvés à travers la procédure d'estimation (des paramètres) sur données réelles. Les résultats en Figure 2 (à gauche) montrent que, si ALCOVE a la plus petite SSD, alors dans 90% des cas les données ont été obtenues en utilisant ALCOVE. Au contraire, si Component-Cue a la plus petite SSD, alors dans 84.8% des cas les données proviennent de Component-Cue. Ces résultats nous permettent d'avoir un regard confiant vis-à-vis des conclusions que nous obtiendrons sur données réelles.

5 Application aux données réelles

Nous pouvons enfin appliquer cette méthode aux données expérimentales recueillies par Mathy [5] pour établir quel modèle, parmi ALCOVE et Component-Cue, se rapproche au mieux aux réponses de chaque individu. Le résultat (Figure 2 à droite), montre que la moitié des participants suit une logique plus proche d'ALCOVE, tandis que l'autre moitié suit une logique plus proche de Component-Cue. Étant donné la fiabilité de l'analyse sur données artificielles, nous pouvons affirmer qu'il y a deux types d'individus : une classe d'individus qui a un raisonnement plus proche d'ALCOVE et une autre qui utilise un raisonnement plus proche de Component-Cue.

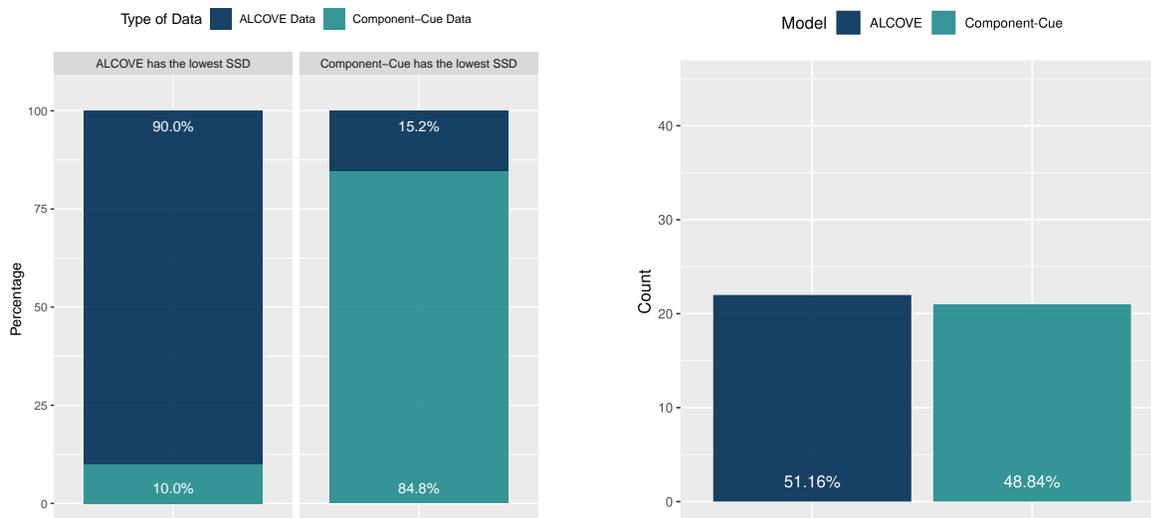


FIGURE 2 – Comparaison entre ALCOVE et Component-Cue. À gauche, sur données simulées, le pourcentage de données simulées provenant d’ALCOVE ou de Component-Cue, en fonction du modèle ayant la plus petite SSD. À droite, sur données réelles, le nombre et le pourcentage d’individus dont les données s’ajoute au mieux à ALCOVE ou à Component-Cue.

Références bibliographiques

- [1] Kruschke, J. K. *ALCOVE : An exemplar-based connectionist model of category learning*. Psychological Review, 1992.
- [2] Gluck M. A. & Bower G. H. *From conditioning to category learning : An adaptive network model*. Journal of Experimental Psychology : General, 1988.
- [3] Nosofsky, R. M. & Kruschke J. K. & McKinley S. C. *Combining Exemplar-Based Category Representations and Connectionist Learning Rules*. Journal of Experimental Psychology : Learning, Memory and Cognition, 1992.
- [4] Nosofsky, R. M. & Gluck, M. A. & Palmeri T. J. & McKinley S. C. & Glauthier P. *Comparing modes of rule-based classification learning : A replication and extension of Shepard, Hovland, and Jenkins (1961)*. Memory & Cognition, 1994.
- [5] Mathy, F., & Feldman, J. *The influence of présentation order on category transfer*. Experimental Psychology, 2016.