

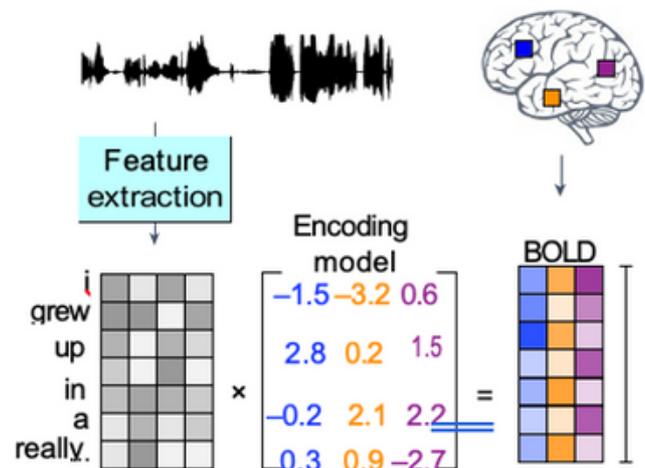
L'IA LANGAGIÈRE AU SERVICE DU « DÉCODAGE CÉRÉBRAL » DE LA PENSÉE

Michael PICHAT, docteur en psychologie des processus cognitifs, maître de conférences des universités, fondateur de neocognition

Fraichement paru dans la prestigieuse revue Nature neuroscience en 2023, un groupe de chercheurs (J. Tang, A. LeBel, S. Jain & A. Huth) en sciences informatiques et neurosciences de l'Université d'Austin (USA), a publié un article intitulé « Reconstruction sémantique du langage continu à partir d'enregistrements cérébraux non invasifs » constituant une avancée majeure dans l'utilisation de l'intelligence artificielle à des fins de « décodage » de la pensée interne humaine.

Fabrication d'une « clé de cryptage sémantique » de la pensée verbale humaine

Les chercheurs nord-américains ont fait écouter à des individus une importante quantité de séquences de mots contenues dans des histoires narratives. Avec une technologie usuelle d'imagerie par résonance magnétique fonctionnelle (IRMf), les scientifiques ont enregistré les variations particulières d'afflux en oxygène sanguin des régions corticales « sémantico-langagières » activées lors de la perception auditive de chacune de ces séquences verbales. Ces enregistrements sont ainsi une image numérique, au sens mathématique du terme, de l'activité neuronale traduisant physiologiquement l'encodage sémantico-langagier de ces séquences verbales.



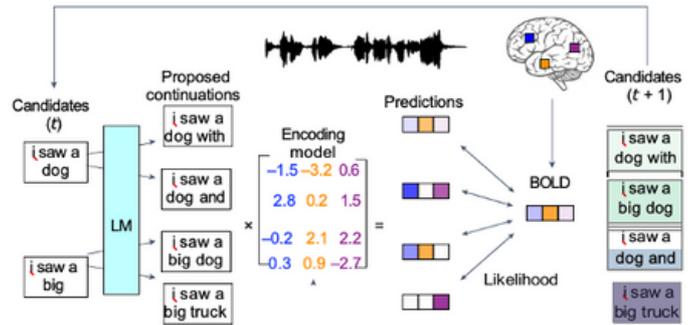
Chaque stimulus de séquence de mots a dès lors été apparié à l'activité électrochimique cérébrale produite en réponse par les individus étudiés, créant de facto un ensemble de couples (séquence verbale en entrée, signature numérique spécifique en sortie). La fonction langagière-numérique ainsi définie représente indirectement les encodages sémantiques réalisés par ces individus. Cette fonction, qui est en fait mathématiquement une matrice de correspondance (input, output) post-traitée par régression, peut être alors interprétée comme une clé de lecture du cryptage sémantique opéré par la cognition humaine (cf schéma ci-après, issu de l'article princeps).

Le « décryptage » des pensées verbales

Si cette fonction de cryptage permet une traduction séquence de mots -> signature neuronale, sa fonction réciproque rend alors théoriquement possible le décryptage, en sens inverse, signature neuronale -> séquence de mots. Et, par extrapolation, le décryptage activité neuronale -> contenu sémantique de la pensée verbale.

Afin de tester cette possibilité vertigineuse, les chercheurs ont dès lors demandé aux individus impliqués dans l'expérience de produire intérieurement (i.e. sans oralisation) de nouveaux discours narratifs. Leurs activités cérébrales spécifiques à cette occasion ont alors été enregistrées à nouveau par IRMf, puis ces enregistrements ont été soumis à la clé de décryptage inverse neuro-sémantique (cf schéma ci-après, issu de l'article).

Cette clé de décryptage n'est, bien entendu, certainement pas bijective (i.e. pas de correspondance univoque terme à terme) ; la procédure ci-dessus génère dès lors une série de candidats de séquences de mots, dont peu sont probables d'un point de vue strictement syntaxique ou sémantique. C'est ici que l'IA générative langagière intervient en rendant possible une évaluation de la plausibilité langagière de la très grande quantité de candidats obtenus ; en opérant une sélection pour ne retenir que le candidat (le contenu inféré de la pensée verbale intérieure produite) le plus probable.



Les pensées verbales ainsi prédites ont été comparées aux pensées verbales effectivement indiquées après-coup par les participants de l'expérience, par mesure de leur proximité sémantique via d'autres techniques d'IA langagière. Les résultats statistiques obtenus indiquent que le sens de ces pensées verbales a été significativement correctement prédit par la clé de décryptage neuro-sémantique arbitrée par l'IA langagière (cf exemples ci-après, issus de l'article).

Avec des garde-fous éthiques draconiens à mettre solidement en place dès maintenant comme pour toute innovation scientifique majeure, les potentielles applications de cette technologie d'utilisation des IA langagières sont vastes et prometteuses ; par exemple dans le champ thérapeutique (citons le cas de l'aphasie de production langagière) et dans celui de l'interface de commande homme-machine. Et, n'en doutons pas, ce domaine de recherche va connaître une croissance aussi forte que surprenante dans les toutes prochaines années.

Actual stimulus	Decoded stimulus	
<i>i got up from the air mattress and pressed my face against the glass of the bedroom window expecting to see eyes staring back at me but instead finding only darkness</i>	<i>i just continued to walk up to the window and open the glass i stood on my toes and peered out i didn't see anything and looked up again i saw nothing</i>	Exact
<i>i didn't know whether to scream cry or run away instead i said leave me alone i don't need your help adam disappeared and i cleaned up alone crying</i>	<i>started to scream and cry and then she just said i told you to leave me alone you can't hurt me anymore i'm sorry and then he stormed off i thought he had left i started to cry</i>	Gist
<i>that night i went upstairs to what had been our bedroom and not knowing what else to do i turned out the lights and lay down on the floor</i>	<i>we got back to my dorm room i had no idea where my bed was i just assumed i would sleep on it but instead i lay down on the floor</i>	Error
<i>i don't have my driver's license yet and i just jumped out right when i needed to and she says well why don't you come back to my house and i'll give you a ride i say ok</i>	<i>she is not ready she has not even started to learn to drive yet i had to push her out of the car i said we will take her home now and she agreed</i>	