

Ouvrir les modèles d'IA, un moyen d'accélérer leur optimisation... et de trouver sa place sur le marché

Mi-octobre, s'est tenue à Paris la conférence dotAI : un événement phare de l'écosystème de l'intelligence artificielle, qui a réuni de nombreuses références du secteur - d'OpenAI à Google Deepmind en passant par HuggingFace ou Kyutai. L'occasion d'évoquer une démarche grâce à laquelle les Français pourraient bien faire la différence : l'ouverture des modèles.

Temps de lecture : minute

21 novembre 2024

C'est l'un des constats qui s'est imposé à l'issue des deux journées de la conférence dotAI, les 17 et 18 octobre, aux Folies Bergère : après une phase marquée par l'essor des modèles fermés - ceux d'OpenAI en tête - des modèles plus ouverts commencent à se faire une place sur le marché de l'IA générative.

“Un marché immense” pour Google

Parmi les récents convertis figure ainsi Google, qui propose depuis février 2024 une série de modèles “ouverts” pour les développeurs - tout en continuant en parallèle à miser sur son modèle propriétaire Gemini. « *Le marché de l'open source est immense. C'est là où se trouve la majorité des développeurs* », a expliqué Armand Joulin, Research Director chez Google Deepmind lors de la conférence. Les modèles Gemma - c'est leur nom - ont ainsi enregistré plus de 20 millions de téléchargements en

moins de six mois.

Pour Google, qui s'est laissé distancer par OpenAI dans la course à l'IA générative, cette démarche d'ouverture est aussi un moyen de rattraper une partie de son retard. Meta, avec ses modèles LLaMA, a d'ailleurs suivi la même stratégie. De son côté, la licorne française MistralAI avait fait dès le départ le choix de l'open source.

Plus d'un million de modèles sur Hugging Face

« L'open source appliqué aux grands modèles permet à la communauté de bâtir au fil du temps de meilleurs modèles, plus petits et plus efficaces. C'est ce qui fait que l'open source finit presque toujours par combler son retard », estime ainsi Merve Noyan, Machine Learning Advocate Engineer chez Hugging Face.

Ce pionnier de l'IA en open source en sait quelque chose : créée en 2016 par les Français Clément Delangue, Julien Chaumond et Thomas Wolf, Hugging Face a su fédérer une communauté très engagée à travers le monde entier, qui fait aujourd'hui toute sa force. La plateforme répertorie par exemple plus de 140 000 modèles de génération de texte, tandis que plus d'un million de modèles sont disponibles pour d'autres tâches, allant de la vision par ordinateur à la reconnaissance automatique de la parole.

La mise à disposition des modèles et des outils en open source permet en effet à une communauté de chercheurs, de développeurs et d'utilisateurs de collaborer, d'expérimenter et d'améliorer les modèles existants, pour les optimiser. L'accès à l'IA est ainsi aujourd'hui largement facilité par l'existence de nombreuses bibliothèques et outils qui simplifient le développement et le déploiement des modèles.

Des données synthétiques pour entraîner Moshi

Pour autant, si les modèles sont “ouverts”, voire totalement open-sourcés, ce n’est pas le cas des données qui servent à les nourrir. « *Les entreprises qui publient des modèles ne publient généralement aucune donnée, car c'est de là que vient l'avantage concurrentiel* », constate Merve Noyan. Mais des solutions communautaires existent, telles que la mutualisation des efforts de collecte et d'annotation - notamment via l’initiative “[Data is Better Together](#)” d’Hugging Face - ou la génération de données nouvelles données, dites “synthétiques”.

C’est d’ailleurs l’approche adoptée par Kyutai, le laboratoire de recherche lancé en novembre 2023 par Xavier Niel, Eric Schmidt et Rodolphe Saadé. Son intelligence artificielle vocale nommée Moshi a été développée grâce à des données “fabriquées” par un système de synthèse vocale capable de cloner des voix et de générer des dialogues, afin de pallier le faible volume d’enregistrements audio disponibles pour entraîner le modèle.

De fait, « *Moshi est le tout premier modèle vocal d'IA en temps réel publié en open source* », souligne Neil Zeghidour, cofondateur de Kyutai et ancien de Google Deepmind. Un choix qui s’est imposé, puisque pour Kyutai, l’« *objectif est de favoriser et de faciliter autant que possible l'adoption de la technologie des agents vocaux à des fins de recherche et de commercialisation* ».

Les chercheurs de Kyutai ont notamment publié un article de recherche de 60 pages contenant tous les détails techniques de Moshi, ce qui permettra à d’autres de comprendre le fonctionnement du modèle et de le reproduire. Pour compléter cette démarche d’ouverture et de transparence, le code d’entraînement du modèle est également en cours de publication.

De quoi permettre à la jeune startup française de s'imposer comme la référence de cette technologie à l'échelle mondiale ? Potentiellement, car, comme le rappelle Merve Noyan, en matière d'IA « aucune entreprise [établie] n'a d'avantage compétitif dans la durée ».



À lire aussi

Après son partenariat avec Microsoft, Mistral AI veut montrer patte blanche vis-à-vis de l'écosystème open source



MADDYNEWS

La newsletter qu'il vous faut pour ne rien rater de l'actualité des startups françaises !

JE M'INSCRIS

Article écrit par Maddyness avec dotConferences