

RABBIT R1

La fin des IHM?

Sommaire

- Notion clé: le LLM..... 1
- Le problème..... 1
- Zoom sur... Mistral7B..... 2
- La fin des IHM?..... 2



Ce minuscule gadget est une nouvelle étape dans l'utilisation de l'IA

Edito

Star du CES 2024, j'ai vu passer ce truc sur tous mes réseaux sociaux, y compris ceux qui concernent la musique car il s'agit d'un partenariat avec Teenage Ingeneering, un fabricant d'instruments électroniques, plus connu pour son design et ses prix exorbitants que pour leur qualité sonore, même si elle est indéniable.

J'ai d'abord pensé que c'était un nouveau gadget à la mode, avant de me résigner à m'y intéresser. J'ai réalisé que je me trompais grandement, sans toutefois avoir complètement tort, même si ça peut paraître paradoxal, nous verrons pourquoi lors de ce hors série spécial Rabbit R1.

Pour les fans de la VT, rassurez vous, la lettre spéciale CES arrive bientôt, celle ci était prête avant, c'est tout.

Loïc Leveau.

Notion clé: le LLM

Un LLM (large language model) est un modèle informatique capable de comprendre et générer du langage naturel. Concrètement, ça sert à avoir une conversation avec Siri, Alexa, pour ceux qui ont une enceinte Amazon, ou chat GPT. Il existe plusieurs modèles, chaque entreprise utilise le sien, et donc, une des batailles, c'est d'avoir le meilleur. On a donc mis en place un classement, réalisé à partir de plusieurs critères, avec des vrais humains qui votent. Vous l'aurez compris, l'efficacité des réponses données par votre appareil va dépendre du LLM qu'il utilise (powered by AI oui, mais laquelle?). En ce moment, c'est le fameux chat GPT qui gagne, mais peut être plus pour longtemps.

Le problème

L'IA est incollable. Elle est très forte pour trouver l'information, et générer du texte, qu'elle va lire à haute voix ou non. Ça tombe bien, c'est ce qu'on lui demande. Mais jusqu'à présent, l'IA ne pouvait pas réaliser une action, sauf dans le cas de Google Home qui était en fait l'appli qui commande votre Smart TV, ou Alexa, qui peut lancer votre playlist sur Amazon music. Parce qu'elle ne pouvait pas ouvrir une autre appli, et cliquer là ou il faut.

Zoom sur... Mistral, l'intelligence artificielle intelligente... et française!

Comme nous l'avons vu, il existe plusieurs modèles de LLM. Un chiffre important à comprendre, c'est le nombre de paramètres. Les paramètres d'un LLM, ça va être, par exemple, sa capacité à faire une traduction dans différents langages, ou à résumer un texte. On pourrait assez vite en déduire que plus il y a de paramètres, et mieux c'est. C'est souvent le cas, je vous rassure, mais il y a un mais.

Plus il y a de paramètres, plus le LLM est compliqué à mettre en œuvre. Un LLM à 70 milliards de paramètres ne va pouvoir être actionné que par des machines ayant des puissances de calcul phénoménales. On va donc les interroger sur un serveur.

Mais imaginons un modèle plus petit, qui aurait moins de paramètres, mais de meilleurs algorithmes. Qui pourrait tourner local sur une machine grand public comme un ordinateur portable. Et imaginons que ce modèle concurrence ChatGPT! Et bien c'est chose faite, et c'est français figurez-vous, ça s'appelle "Mistral". Ça implique donc que des petites entreprises peuvent héberger un LLM sur leur propre serveur. Et ça implique donc que la société française Mistral devient un concurrent sérieux pour des géants américains, en proposant une solution à la portée de tout le monde. Plus besoin des serveurs de Google pour proposer une IA!

Le Rabbit R1, la fin des IHM?

C'est à ce moment là que le Rabbit R1 entre en action. Apparemment, un smartphone, c'est encore trop compliqué à utiliser, il faut un ordinateur qui l'utilise à votre place. Pfiou! Moi évidemment, avec mes 3 applis, Waze, YouTube, et la météo, ça me fait un peu rigoler, mais je ne devrais évidemment pas, puisque la techno est prometteuse, même si ce boîtier ne va atterrir que dans très peu de poches, c'est quand même une étape de plus dans l'utilisation de l'IA.

Le Rabbit R1 est une interface entre vous et toutes vos applications. Au lieu de cliquer sur vos applis dans votre smartphone, vous parlez à votre Rabbit. Ça peut faciliter la vie, mais c'est surtout une étape dans l'utilisation de l'IA aux mains du grand public. En gros, vous avez une Alexa entre vous et votre smartphone. Vous pouvez lui demander de vous jouer tel morceau de musique, de vous réserver une nuit d'hôtel, un billet de train, commander un repas, que sais-je encore. Je ne suis pas le mieux placé pour vous donner des exemples d'utilisation, n'étant pas moi un grand fan des smartphones. Je suis tout simplement ce qu'on peut appeler un boomer.

Cette machine sait utiliser vos applis les plus courantes, et vous pouvez même lui apprendre à en utiliser de nouvelles. Par exemple, vous pouvez dire à votre Rabbit: "ouvre Waze et trouve moi le meilleur itinéraire pour rentrer à la maison", ce qui, au volant d'une voiture, est, il faut bien le reconnaître, incroyablement pratique, et vous pouvez ajouter: "et commande-moi des sushis pour 20 H". D'après ses créateurs, vous pouvez même lui donner des ordres tels que: "prend moi un billet d'avion pour New York le 3 février, une voiture de location, et réserve moi un hôtel à Manhattan pour la nuit, à 10 minutes à pied de Central Park, et dans une rue animée avec des bars branchés." Bon, il faut avoir la maxi confiance en son appareil, et énormément de sous, mais, c'est ça l'idée. Si vous utilisez une appli particulière pour traiter vos photos, vous pouvez même lui montrer comment vous faites, puis elle saura le faire toute seule. Vous donc alors passer l'ordre suivant: "retouche toutes les photos que j'ai prises hier et avant hier et met les en ligne sur mon insta."

Ça va plus loin que Siri ou Alexa. Ça fait des actions. Ça clique à votre place. Voilà pourquoi les créateurs parlent de remplacer les LLM (large language model) par un LAM, pour Large Action Model.

Alors est-ce que c'est une révolution? Oui et non. Dans les faits: ils vont en vendre plein, se faire plein de sous, et terminé. Les requêtes vont devenir délirantes et la boîte va fermer, parce que c'est vendu au prix de 200 dollars mais qu'il n'y a pas d'abonnement. Si je vous en parle, c'est juste parce que la technologie est intéressante. Si je peux me permettre une petite prévision, comme ça, au débotté, ce qu'il va se passer c'est que le concept sera incorporé assez vite dans les smartphones les plus luxueux, tout simplement. Et, vraisemblablement, dans des machines pour professionnels. On peut imaginer un chirurgien donner des ordres oraux, dans un futur relativement proche, à une machine plus précise et plus rapide que lui...