

Les chiffres ahurissants dévoilés : combien coûte le fonctionnement d'un chatbot ?

ChatGPT attire de plus en plus d'utilisateurs mais coûte cher à OpenAI

ChatGPT, le chatbot développé par OpenAI, a connu une croissance rapide en termes d'utilisateurs depuis son lancement en novembre dernier, comptant aujourd'hui plus de quatre milliards de visites et 186 millions de comptes en mars. Cependant, son fonctionnement nécessite des coûts considérables, notamment pour assurer une maintenance extrêmement coûteuse pour maintenir le chatbot opérationnel. Selon une estimation révélée par l'analyste en chef chez SemiAnalysis, Dylan Patel, le chatbot nécessiterait 700 000 dollars par jour pour rester opérationnel, soit 36 cents par requête. La majeure partie de ce coût est basée sur les serveurs coûteux dont ChatGPT a besoin.

Les coûts d'inférence dépassent les coûts de formation d'entraînement

Bien que la formation des grands modèles linguistiques de ChatGPT coûte probablement des dizaines de millions de dollars, les coûts d'inférence – c'est-à-dire les coûts d'exécution des modèles sur des données clients – dépassent de loin les coûts de formation d'entraînement lors du déploiement d'un modèle à une telle échelle, expliquent Patel et Afzal Ahmad, un autre analyste de SemiAnalysis. En fin de compte, les coûts d'inférence ChatGPT dépassent les coûts de formation sur une base hebdomadaire.

Microsoft travaille sur une puce pour aider à réduire les

coûts

Afin de réduire les coûts, Microsoft développe une puce connue sous le nom d'Athena pour aider au fonctionnement quotidien de ChatGPT. En chantier depuis 2019, celle-ci permettrait de réduire considérablement les coûts de fonctionnement du robot conversationnel développé par OpenAI. Plus de 300 employés de Microsoft travailleraient désormais activement sur son développement. La puce pourrait être prête pour un usage interne par Microsoft et OpenAI dès l'année prochaine.

Conclusion

ChatGPT connaît un succès indéniable auprès des utilisateurs, mais son fonctionnement est coûteux pour OpenAI. Le chatbot nécessiterait 700 000 dollars par jour pour rester opérationnel, soit 36 cents par requête en raison de l'infrastructure coûteuse sur laquelle il repose. Une puce développée par Microsoft pourrait aider à réduire considérablement ces coûts à l'avenir.

Source : tomsguide.fr

→ **Accéder à [CHAT GPT](#) en cliquant dessus**