

Comparaison des deux modèles

Modèle générique :

Pour le modèle générique, nous avons utilisé la bibliothèque python transformers pour permettre de rapidement importer GPT-2 et pouvoir le tester des des entrées quelconques.

Etant donné que le modèle n'était pas spécialement entraîné sur le domaine de la finance mais plus sur une large banque de données tel que des livres et écrit de tout type, les réponses n'étaient pas forcément pertinentes.

Exemple avec comme entrée : "What is the cost of most tech companies", nous obtenons la réponse suivante :

"What is the cost of most tech companies?"

"The costs of the companies in the U.S. are often higher than what we are paying for those companies," he said. "I guess it's kind of one of those things where the cost is going down because they're making a lot of money, but I think this is one of the more complex and complicated parts of our economy."

Read more:

In 2014, the average American household invested \$4,8"

On voit déjà un problème, c'est que le modèle GPT-2 ne fait que compléter l'entrée qu'on lui donne avec des données initiales qui n'ont pas spécialement de sens.

Modèle entraîné :

Pour pouvoir entraîner le modèle sur le dataset de finance de hugging face, nous sommes partis de l'API keras-nlp comme notre professeur nous a conseillé. A la fin de l'entraînement avec 30 epochs, qui a durée environ 45 minutes, nous avons exporté le modèle sous forme d'un fichier de 1.5 GB pour éviter de devoir le réentraîner à chaque utilisation.

Nous allons donc lui donner la même entrée et voir si sa réponse est plus pertinente.

Entrée : "What is the cost of most tech companies"

Réponse :

"What is the cost of most tech companies building an infrastructure project in India ?"

On peut remarquer qu'il est déjà plus pertinent et nous donne bien la question de manière plus complète et précise.

Entrée : "The stock market is" → "The stock market is bullish about Qualcomm 's product development and about the potential of the new platform ." Ici, nous avons une réponse plus précise en citant une marque d'électronique.

Entrée : "Why the amount of capital in companies are low ?" → "Because the companies are too big to manage ." Avec ce dernier exemple, le modèle complète la phrase en nous donnant une réponse très simple. C'est le problème de ce dataset qui n'est pas assez riche.

Remarques :

Nous pouvons voir que le principal problème de GPT-2 à compléter uniquement l'entrée de l'utilisateur est aussi présent malgré l'entraînement. Avec le modèle entraîné nous remarquons que les réponses sont plus précises et se base sur des cas concrets de finance correspondant à notre dataset.

Cependant, nous pensons que le vrai problème de notre modèle est le dataset associé qui est très limité, car en intelligence artificielle le plus important est le volume et la qualité des données.

Il y a aussi une chose à améliorer, c'est le temps de réponse du modèle qui pour des entrées assez précises peut aller de 5s à 30s.