

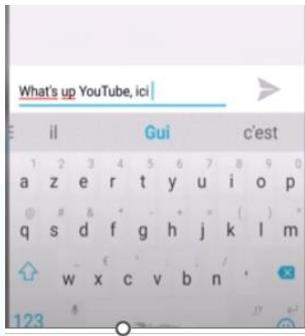
L'AGENT CONVERSATIONNEL : CHATGPT

Introduction

L'un des développements les plus remarquables de l'IA à l'heure actuelle est l'émergence de l'agent conversationnel **ChatGPT** créée par l'entreprise **Open AI**. Mais, avant d'expliquer son principe de fonctionnement, nous clarifions tout d'abord deux concepts différents, souvent confondus : **GPT** et **ChatGPT**

1. GPT vs. ChatGPT

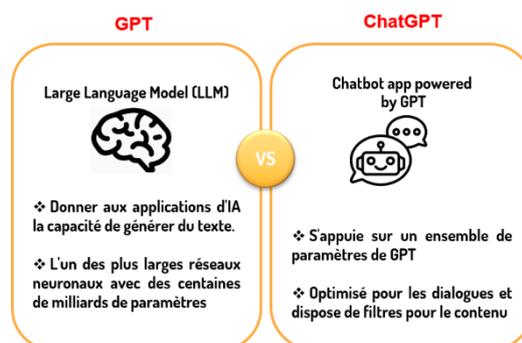
GPT (Generative Pre-trained Transformer) est un langage de type LLM (**Large Language Model**) pour le traitement et la génération du langage naturel. Ces modèles sont des modèles d'apprentissage automatique qui sont extrêmement efficaces lorsqu'il s'agit d'effectuer des tâches liées au langage : traduire, résumer des textes, générer du contenu ou du code, etc.



GPT utilise l'écriture prédictive du LLM pour donner des réponses cohérentes en utilisant la suite de mots la plus probable selon le contexte. **Eh ben, avec votre mobile, vous utilisez quotidiennement le langage LMM dont se sert ChatGPT !**

Dans nos applications de messages ou moteurs de recherche, nous utilisons souvent le texte intuitif qui se base généralement sur un dictionnaire préexistant et des probabilités.

Quant au concept **ChatGPT** : C'est un mot-valise : "chat" qui fait référence à **une discussion en ligne** et "GPT" qui signifie que ChatGPT a été **préalablement entraîné** pour **générer** des réponses pertinentes selon le **contexte**.

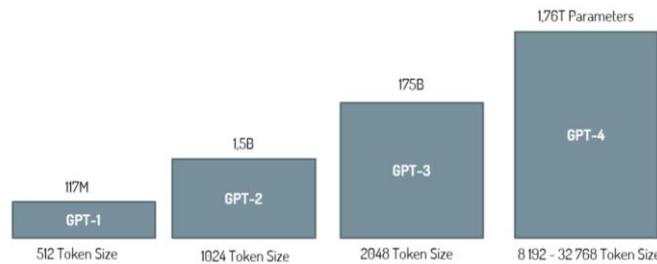


Ce qu'il faut retenir c'est que **GPT** est un modèle de langage qui sert de « **cerveau** » à l'agent conversationnel **ChatGPT** et ce, pour tenir des conversations fluides et naturelles avec les utilisateurs. Nous expliquons par la suite les différents stades de l'évolution des modèles GPT.

2. Les stades d'évolution des modèles GPT

Le concept GPT a été initialement **introduit** en juin 2018 avec la publication du premier modèle, **GPT-1**, puis les autres modèles **GPT-2**, **3** et **4** sont dévoilés en **2019**, **2020** et **2023 respectivement**. La principale différence réside dans la taille ainsi que les paramètres qui sont

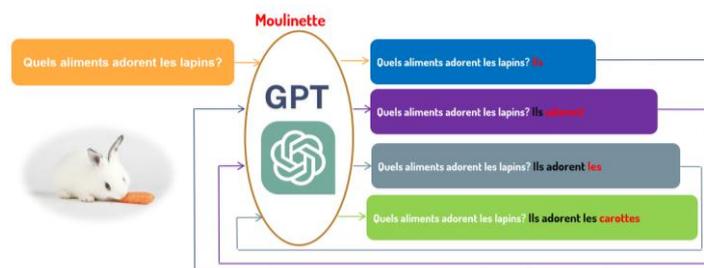
passés de 117 millions dans GPT-1 à 1.76 Trillion dans GPT-4. D'autant plus, GPT-4 peut recevoir maintenant des images ce qui constitue une amélioration majeure puisque les versions précédentes de GPT acceptent uniquement les inputs sous forme de texte.



3. Fonctionnement de ChatGPT

Quand on fournit à ChatGPT un « **prompt** » (que ce soit **une question ou des instructions**), Il prédit dans un premier temps le mot le plus probable qui doit suivre en s'appuyant sur des millions de textes que le modèle a intégrés lors de **son apprentissage**.

Ensuite, il ajoute ce mot au prompt d'origine et renvoie l'ensemble à la case d'entrée pour le repasser à la moulinette (GPT). Puis, un 2^{ème} mot est ajouté et ainsi de suite jusqu'à un caractère spécial indique la fin de la réponse.



GPT essaie de produire une suite de mots qui semble appropriée dans le contexte donné.

- **GPT analyse le contexte et génère une réponse en utilisant les connaissances acquises pendant son apprentissage.**
- Il faut noter que CHATGPT est affiné grâce à l'utilisation des techniques d'apprentissage supervisé et par renforcement du modèle GPT. Mais comment ?

Nous expliquerons les différents modes d'apprentissage de GPT que ce soit à l'aide :

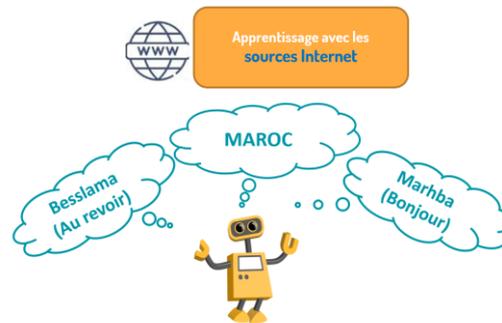
- Des Sources d'Internet
- De la Supervision Humaine
- La Supervision de IA

4. Les modes d'apprentissage de GPT

Le mode d'apprentissage avec les sources d'internet.

Google a publié son article « **Attention Is All You Need** » en 2017 et qui a fait un grand pas en avant dans le domaine avec l'avènement de la technologie des « **Transformers** ». ChatGPT utilise cette technologie, **pour capturer les relations entre les mots** dans un texte en se basant sur de vastes quantités de **données textuelles provenant d'Internet**, articles, journaux, livres, réseaux sociaux, ...etc.

Par exemple, en analysant l'expression suivante : « Lors de votre visite au Maroc, les salutations les plus communes que vous entendrez seront Marhba (bonjour) et Besslama (au revoir) ». Notre attention sera focalisée sur certains mots qui constituent le cœur de cette expression (ou le contexte), et qui sont : **Maroc, Marhba et Besslama** car notre expérience nous a déjà appris que c'est là que se trouve l'essentiel de l'information.



Lors de votre visite au **Maroc**, les salutations les plus communes que vous entendrez seront **Marhba** (bonjour) et **Besslama** (au revoir)

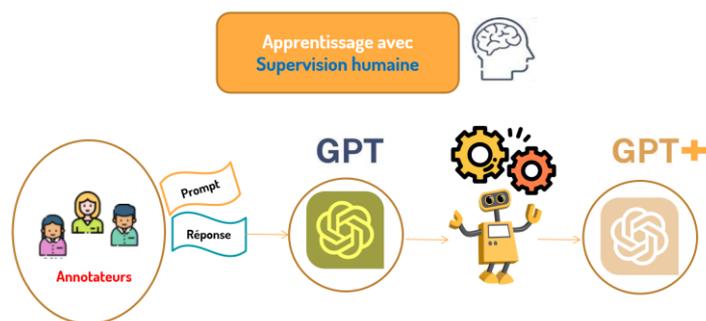
17

Eh bien, les « **Transformers** » utilisent une simulation de ce mécanisme qu'on appelle le « **mécanisme d'attention** ». Il consiste à pondérer (donner un poids) les **mots** d'une phrase qui font partie au **même contexte** afin de les valoriser.

En général, les modèles de langages LLM se basent sur ces architectures de réseaux de neurones profonds « **Transformers** » pour une meilleure compréhension du **contexte** mais aussi pour une **meilleure prise en charge de longues phrases**.

Le mode d'apprentissage avec supervision humaine : Ce modèle comprend deux méthodes:

- La première méthode du supervision humaine fait appel à des annotateurs : Des personnes qui rédigent un **ensemble de questions (prompts) et leurs réponses**

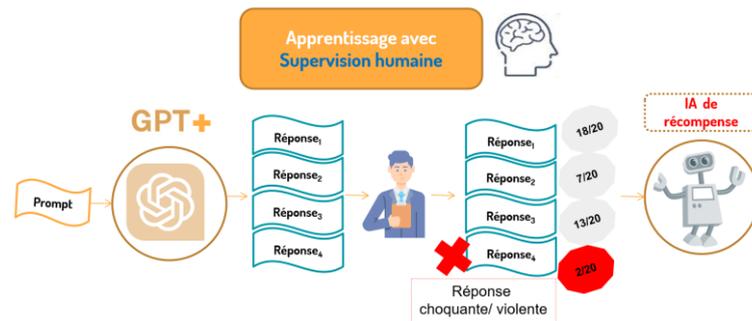


Supervision Humaine à l'aide des annotateurs

19

GPT est alors entraînée en se basant sur ces données (prompts/réponses) fournis par les annotateurs jusqu'à ce qu'elle donne **la réponse attendue à la demande (prompt) initiale**. On obtient ainsi une version améliorée de GPT → (notée ici GPT+)

- La deuxième méthode avec supervision humaine sera basée ensuite sur un **modèle conversationnel** appelé : **RLHF** (« Reinforcement Learning from Human Feedback »).



Supervision Humaine à l'aide de RLHF*

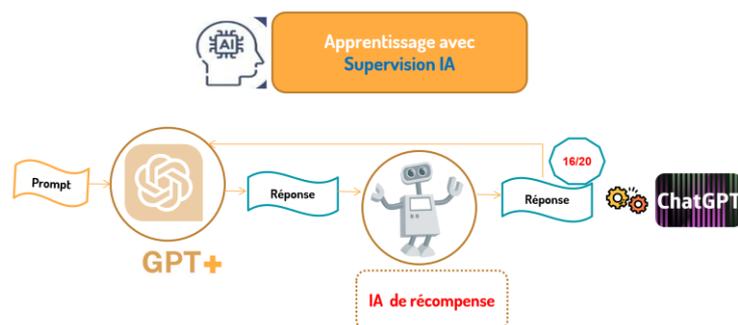
* Reinforcement Learning from Human Feedback

20

Son principe est simple : On demande à GPT de donner **plusieurs réponses à une même question (prompt)**. Puis, des superviseurs classent et évaluent la pertinence de ces **réponses** en donnant à chaque réponse une “récompense” (“rewards”). Les superviseurs pointent aussi les réponses **choquantes ou violentes** afin de renforcer la cohérence des réponses et éviter les dérives et les erreurs. Toutes ces données (**réponses et notes**) serviront pour entraîner une nouvelle IA (IA de récompense) à comprendre par la suite comment classer et évaluer les réponses.

Le mode d'apprentissage avec l'IA de récompense.

Durant la dernière phase d'apprentissage, de nouvelles questions sont données à GPT+ **sans supervision humaine**. Chaque réponse sera évaluée cette fois-ci par **l'IA de récompense**. Plus la note est haute, plus GPT+ apprend à valoriser la réponse. Sinon, la réponse sera marginalisée. A la fin de l'apprentissage, on obtient la version actuelle de : ChatGPT.



Supervision à l'aide de l'IA de récompense

21

Conclusion

Pour conclure, ChatGPT fonctionne **de manière interactive** en échangeant des tours de parole avec un utilisateur selon le modèle GPT. D'autant plus, CHATGPT est affiné grâce à l'utilisation des « Transformers » et des techniques d'apprentissage supervisé et par renforcement du modèle GPT.

Parmi toutes les réponses linguistiquement correctes que pourrait générer, CHATGPT opte pour celles qui sont le plus correctes. Sur la base de ce même principe, CHATGPT refuse aussi parfois carrément de répondre.



Pr. CHARAF MOULAY EL HASSAN
Université IBN TOFAIL- Kenitra