TD 1 — Bases IAg et Prompting

1. Quelques manipulations simples avec Vittascience

Nous allons utiliser ici l'outil Vittascience[1].

Exercice 1 — Compréhension des IAg textuelles 1

L'objectif de cet exercice est de comprendre le fonctionnement d'un modèle de langage textuel à travers un agent conversationnel.

Pour cela :

1. connectez-vous sur Vittascience : https://fr.vittascience.com/ia/text
2. Positionnez le curseur « Aléatoire » à 50%.
3. Choisir le modèle « Mistral Small 3 »
4. Entrer la requête : « Qu'est-ce qu'un chat ? »
5. Faire le « Pas à pas »
6. Cliquer sur « Montrer » pour visualiser les tokens.
7. Cliquer sur quelques tokens pour voir leur probabilité et les autres possibilités.
8. Cliquer sur « Token ID »
9. Cliquer sur « Voir les vecteurs de mots »
10. Revenez sur les mots et choisissez un token différent, où vous le souhaitez, dans la réponse.
11. Recommencer en mettant « Aléatoire » à 5% puis à 90%

> **Remarque**
>
> Attention, avec Vittascience, le quota est de 30 requêtes par mois, 10 par jour. Il faut donc être économe !

Exercice 2 — Compréhension des IAg textuelles 2

Reprendre l'exercice précédent, mais avec le prompt :
« Qu'est-ce que l'acide désoxyribonucléique ? »

Exercice 3 — Compréhension des IAg textuelles 3

L'objectif de cet exercice est de travailler un peu les prédictions.

Proposer des suites possibles à « Indique la suite de : il était une fois un lapin blanc… ». Tester sur un agent conversationnel. Ajouter, dans le prompt, des indications pour le style de ce qui doit être produit.

[1] En cas de souci (dépassement du quota par exemple) vous pouvez vous rabattre sur https://asql.ls2n.fr/rapidchat/

Exercice 4 — Compréhension des IAg d'images

L'objectif de cet exercice est de comprendre le fonctionnement d'un modèle de langage permettant de générer des images.

Pour cela :

1. connectez-vous sur Vittascience : https://fr.vittascience.com/ia/image-generate
2. Positionnez le curseur « Aléatoire » à 50%.
3. Choisir le modèle « stable-diffusion-xl »
4. Entrer la requête : « Une tortue terrestre faisant du surf. » (Prompt) et « Pas de tortue marine. » (Negative prompt).
5. Faire « Générer »
6. Cliquer sur « Etapes » pour visualiser les différentes étapes de génération.
7. Recommencer en mettant « Guidance » à 15

Remarque

Attention, avec Vittascience, le quota est de 10 images par mois, 3 par jour. Il faut donc être très économe !

Exercice 5 — Un texte fluide est-il un texte vérifié ?

*L'objectif de cet exercice est de constater qu'un modèle produit une réponse **plausible**, pas une réponse **vérifiée**. Durée : environ 15 minutes.*

1. Choisissez un ouvrage classique peu célèbre, disponible gratuitement sur Internet, et demandez à 2 ou 3 modèles (par exemple Duck.ai, avec et sans recherche web) la **première phrase** de cet ouvrage.
2. Vérifiez chaque réponse sur le texte original.
3. Signalez l'erreur au modèle : que répond-il ? corrige-t-il vraiment sa réponse ? quelle confiance affiche-t-il ?
4. Notez vos observations (réponse, confiance affichée, réaction à la correction), puis reformulez votre constat **sans** verbe d'intention (« ment », « sait », « veut »...).

Remarque

Ne choisissez pas une première phrase trop célèbre : le modèle l'a probablement mémorisée. Les résultats varient selon le modèle et sa version : comparez-les avec ceux de vos camarades.

2. Agent conversationnel

Pour cette partie, nous allons travailler avec Duck.ai et AnythingLLM du cours.

Exercice 6 — Analyse de texte

Effectuer des traitements sur un texte.

Récupérer le texte : Codd, E.F. (1970). A relational model of data for large shared data banks. Communications of the ACM, 13, 377 - 387. https://www.seas.upenn.edu/~zives/03f/cis550/codd.pdf

Objectifs du travail - Proposer un prompt pour :

- Préciser le contexte (article scientifique).
- Formuler clairement l'objectif (résumé fidèle et structuré).
- Mentionner l'importance de rester fidèle à l'auteur.
- Préciser la longueur attendue (une page maximum).

Dans un second temps :

- Comparer avec le résumé proposé par l'auteur
- Changer de modèle de langage (SLM et LLM) et comparer avec les résumés précédents
- Si c'est possible, modifier la température, le top-p et top-k.
- Formater le résultat : limiter efficacement la taille avec un paramètre adapté et proposer une version structurée.
- Expliciter la cible du résultat : chercheur, étudiant, lycéen, enfant de CP...

Puis :

- Quels sont les opérateurs qu'il présente ?
- Comparer ce qu'il propose avec l'algèbre relationel actuel.
- Quel est l'impact de ce travail ?

Remarque

En réalité, même à température 0, tout n'est pas toujours parfaitement déterministe. Même si l'algorithme est théoriquement déterministe (greedy), sur GPU, de minuscules variations d'arrondi peuvent apparaître car les opérations flottantes ne sont pas parfaitement associatives. Ces écarts peuvent, rarement, changer le token sélectionné.

Ce document est distribué selon les termes de la licence **Creative Commons BY-ND 4.0**. Vous pouvez trouver le texte de la licence à l'adresse suivante : https://creativecommons.org/licenses/by-nd/4.0/

Il utilise la police **Atkinson Hyperlegible Next** (Braille Institute of America, Inc.) afin d'améliorer la lisibilité.

ASSISTÉ PAR L'IA