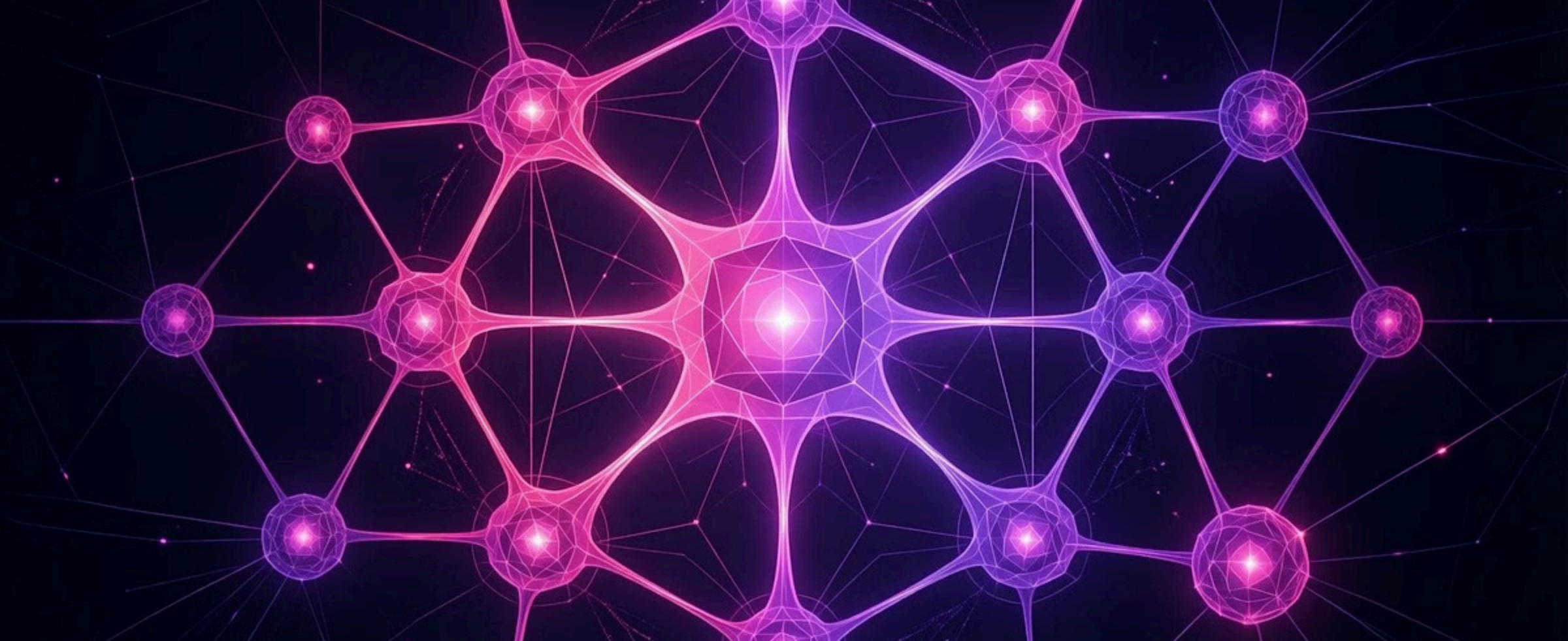




Les LLM

Comprendre comment fonctionnent
les IA qui génèrent du texte

Swipe pour apprendre →



INTRODUCTION

LLM : de quoi parle-t-on ?



Large Language Model

Un modèle statistique entraîné sur un volume énorme de textes.



Objectif principal

Le modèle prédit le prochain élément de texte le plus probable, puis répète ce mécanisme étape par étape jusqu'à produire une réponse complète.



La réponse la plus probable

CONCEPTS CLÉS

Il ne vérifie pas la vérité

Il sait seulement si une suite de mots est probable.

Il ne cherche pas l'information

Contrairement à un moteur de recherche, il ne retrouve rien.

Il construit une réponse

Il génère statistiquement un texte plausible, sans obligation de vérité.

Que se passe-t-il quand on écrit un prompt ?

01

Découpage en tokens

Le texte fourni est découpé en tokens, en fonction de la langue et de ce que le modèle a appris pendant son entraînement.

02

Transformation en nombres

Chaque token est transformé en nombres, afin de pouvoir être traité par le modèle.

03

Analyse du contexte

Le modèle analyse l'ensemble du contexte en prenant en compte tous les tokens précédents.

04

Calcul des probabilités

Il calcule une probabilité pour chaque token possible.

05

Sélection et répétition

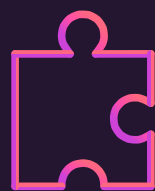
Il sélectionne un token selon les probabilités calculées, puis répète ce mécanisme token par token jusqu'à produire une réponse complète.

C'est quoi un token ?

Un LLM ne manipule pas directement des phrases complètes, il travaille avec des unités plus petites appelées **tokens**.



Un mot entier ou une partie de mot



Un préfixe ou un suffixe



Un signe de ponctuation ou même un espace

L'hallucination des LLM



Qu'est-ce que c'est ?

Une conséquence naturelle du fonctionnement prédictif des LLM,
et non un bug du système.

2

Les LLM sont entraînés à produire une réponse cohérente,
même lorsque les informations sont incomplètes ou absentes.

Par défaut, ils ne répondent pas “je ne sais pas”.



Comment s'en protéger ?

- Prompts rigoureux, mais pas **une solution miracle.**
- Règles et protocoles de validation.
- Vérification humaine des sources.

À RETENIR

Points essentiels sur les LLM



Modèle statistique

Il prédit le prochain élément de texte.



Pas de compréhension

Il imite le langage sans comprendre.



Hallucinations normales

C'est une conséquence normale de son fonctionnement, pas un bug



Encadrement nécessaire

Usage encadré et sources maîtrisées

Merci de votre attention !

J'espère que ce post vous a éclairé sur le fonctionnement des LLM.

N'hésitez pas à le partager et à me suivre pour découvrir les prochains contenus.

Sarah LEON