Comment une IA écrit : le cours complet, zone 3 Expert

Tout ce que le jeu explique, station par station, à lire d'une traite. Zone 3, niveau expert.

Cours complet de La ligne IA · laligne-ia.fr · version expert (Zone 3) · mis à jour le 20 novembre 2026. Le jeu gratuit pour comprendre comment marche une IA.

1 Ligne 1 : Les bases, sous le capot

Découper : Couper le texte en morceaux

Une IA ne lit pas lettre par lettre. Elle coupe d'abord le texte en petits morceaux appelés **tokens** : des mots entiers, ou des bouts de mots qui reviennent souvent.

Le découpage est appris avant tout le reste : on part des lettres et on colle ensemble les paires qui reviennent le plus souvent, des dizaines de milliers de fois. On obtient un vocabulaire de 50 000 à 200 000 morceaux. Effet secondaire : une IA ne « voit » pas les lettres, seulement des morceaux. C'est pour ça que les IA ont longtemps eu du mal à compter les « r » de « carrefour ». Autre détail : le français coûte un peu plus de tokens que l'anglais, parce que le découpage a surtout été appris sur de l'anglais.

> **Sous le capot :** Le découpage est fait par un algorithme de type BPE (byte pair encoding), qui fusionne au départ les paires de caractères les plus fréquentes du corpus. Les modèles récents ont un vocabulaire de 100 000 à 200 000 tokens, et un mot français prend souvent plus d'un token.

> **À retenir : Pour une IA, un texte est une suite de morceaux (les tokens), chacun avec un numéro.**

Lire l'article complet : C'est quoi un token en IA ?

Ranger : Ranger les mots sur une carte

Chaque morceau reçoit une place sur une immense carte. Les mots qu'on utilise dans les mêmes phrases finissent voisins. C'est comme ça que l'IA « sait » que deux mots ont un sens proche.

La vraie carte n'a pas 2 dimensions mais plusieurs milliers : impossible à dessiner, mais le principe est le même, une distance entre les mots. Personne ne place les mots à la main : leur position se règle toute seule pendant l'entraînement. Et elle s'ajuste au contexte : « avocat » ne se range pas pareil dans « mon avocat plaide » et dans « un avocat bien mûr ».

> **Sous le capot :** Chaque token devient un vecteur d'embedding : une liste de plusieurs milliers de nombres. Deux sens proches donnent deux vecteurs proches, ce qu'on mesure par exemple avec la similarité cosinus.

> **À retenir : Pour une IA, le sens d'un mot, c'est sa place sur la carte : des mots proches ont des sens proches.**

Lire l'article complet : Comment une IA comprend-elle le sens des mots ?

Relier : Relier les mots entre eux

Dans une phrase, les mots dépendent les uns des autres. Pour comprendre « elle », il faut savoir de qui on parle. L'IA calcule, pour chaque mot, quels autres mots l'aident le plus. On appelle ça l'**attention**.

L'IA fait ce calcul avec plusieurs « têtes » d'attention en parallèle : l'une suit les pronoms, une autre les accords, une autre les guillemets… Puis elle recommence sur des dizaines de couches empilées, chacune affinant la compréhension. Cette architecture s'appelle le **Transformer**. Elle a été inventée en 2017 par des chercheurs de Google, et c'est le T de GPT.

> **Sous le capot :** Le mécanisme d'attention calcule pour chaque token un score avec tous les autres, à partir de trois vecteurs : requête, clé et valeur. Un Transformer empile des dizaines de couches, chacune avec plusieurs « têtes » qui repèrent des relations différentes.

> **À retenir : L'attention permet à l'IA de relier chaque mot à ceux qui lui donnent son sens.**

Lire l'article complet : L'attention, le secret des Transformers

Deviner : Deviner le mot suivant

Tout ce travail sert à une seule chose : deviner le mot suivant. L'IA ne donne pas une seule réponse, elle donne une chance (en %) à chaque mot possible.

À chaque mot, l'IA donne une probabilité à **chacun** des morceaux de son vocabulaire, soit environ 100 000 notes à chaque fois. Une réponse de 500 mots, c'est 500 fois ce calcul complet. Ça paraît bête, « deviner le mot suivant », mais pour bien deviner la fin d'une démonstration de maths ou d'une recette, il faut avoir compris les maths ou la cuisine.

> **Sous le capot :** À chaque étape, le modèle donne une probabilité à chaque token du vocabulaire (fonction softmax), en choisit un, l'ajoute au texte et recommence. On parle de génération autorégressive : la réponse s'écrit token par token.

> **À retenir : Au fond, une IA de ce type est une machine à deviner le mot suivant, extrêmement bien entraînée.**

Lire l'article complet : Comment une IA comme ChatGPT écrit-elle ses réponses ?

Écrire : Écrire mot après mot

Pour répondre, l'IA écrit un morceau, puis le suivant, puis le suivant… À chaque fois, elle tire au sort parmi les mots possibles, en favorisant les plus probables. Un réglage, la **température**, décide de la dose de hasard.

Pourquoi ne pas toujours prendre le mot le plus probable ? Parce que le texte devient plat et répétitif, et qu'il tourne parfois en boucle. Un peu de hasard rend l'écriture plus naturelle. D'autres réglages existent, comme ne tirer que parmi les mots les plus probables pour éviter les absurdités. Pour du code ou un calcul, on met peu de hasard ; pour un poème, un peu plus.

> **Sous le capot :** La température divise les scores avant le softmax. En dessous de 1, les choix probables écrasent les autres ; au-dessus, la distribution s'aplatit. On la combine souvent avec le top-p, qui ne garde que les tokens couvrant par exemple 90 % de la probabilité.

> **À retenir : Peu de hasard : des phrases sages, toujours les mêmes. Beaucoup de hasard : des phrases originales… puis absurdes.**

Lire l'article complet : La température d'une IA, c'est quoi ?

2 Ligne 2 : Fabriquer et faire tourner

Apprendre : Apprendre de ses erreurs

Au début, l'IA devine au hasard. On lui montre des phrases réelles : elle devine, on compare avec le vrai mot, et on la corrige un tout petit peu. On répète ça des milliards de fois.

Chaque correction modifie un peu les centaines de milliards de réglages internes (les **poids**) en même temps, grâce à une méthode de calcul appelée rétropropagation. On l'entraîne sur des milliers de milliards de morceaux de texte : pages web, livres, code… Un humain qui lirait jour et nuit mettrait des dizaines de milliers d'années pour en lire autant. Personne n'écrit de règle de grammaire : elle émerge parce qu'elle aide à deviner juste.

> **Sous le capot :** L'entraînement minimise une fonction de perte, l'erreur de prédiction du token suivant, par descente de gradient. La rétropropagation calcule dans quel sens ajuster chacun des milliards de paramètres, lot après lot.

> **À retenir : Apprendre, pour une IA, c'est se tromper, être corrigée un peu, et recommencer des milliards de fois. Elle n'apprend pas les phrases par cœur : elle en tire des règles.**

Lire l'article complet : Comment entraîne-t-on une IA ?

Éduquer : Apprendre à bien répondre

Une IA qui devine la suite d'un texte n'est pas encore un bon assistant : elle complète, elle ne répond pas. Des personnes comparent deux de ses réponses et choisissent la meilleure. L'IA est ensuite poussée vers ce genre de réponse.

Avec des milliers de comparaisons, on entraîne un second modèle, le « juge », qui apprend à prédire la réponse que les gens préféreraient. L'IA s'entraîne ensuite à plaire à ce juge. Le risque est connu : à force de vouloir plaire, une IA devient flatteuse et approuve tout. Pour l'éviter, on lui apprend aussi des principes. Chez Anthropic, ils sont réunis dans un texte appelé la « constitution ».

> **Sous le capot :** Après le pré-entraînement viennent le réglage fin sur des exemples de dialogues, puis l'apprentissage par renforcement à partir de préférences humaines (RLHF) ou de principes écrits (IA constitutionnelle).

> **À retenir : Cette étape apprend à l'IA à être utile et honnête, plutôt qu'à flatter ou à refuser pour rien.**

Lire l'article complet : Comment apprend-on à une IA à bien répondre ?

La machine : Le voyage d'un message

Une IA, ce n'est pas un nuage magique. Ce sont des milliers de puces spéciales, rangées dans d'immenses centres de données, qui font des milliards de multiplications par seconde pour calculer chaque mot. Un grand centre de données consomme autant d'électricité qu'une ville.

Ces puces (GPU, TPU) sont faites pour multiplier d'énormes tableaux de nombres en parallèle. Souvent, le plus lent n'est pas le calcul mais le transfert des poids entre la mémoire et le calcul. Pendant une conversation, les calculs déjà faits sur le début du texte sont gardés de côté pour ne pas tout refaire à chaque mot. C'est pour ça qu'une très longue conversation coûte plus cher. Et l'entraînement, lui, mobilise des dizaines de milliers de puces pendant des mois, sur un seul calcul géant.

> **Sous le capot :** Un grand modèle s'entraîne pendant des semaines sur des milliers de GPU reliés en réseau. Pour répondre, ses paramètres doivent tenir dans une mémoire très rapide (HBM) : c'est souvent la mémoire, plus que le calcul, qui limite la vitesse.

> **À retenir : Derrière chaque réponse, il y a des puces spécialisées, beaucoup d'électricité, et un calcul refait pour chaque mot.**

Lire l'article complet : Sur quelles machines tourne une IA ?

L'empreinte : Ce que coûte une réponse

Chaque réponse d'IA consomme de l'**électricité**, et souvent de l'**eau** pour refroidir les machines. Une question isolée coûte peu. Mais des milliards de questions par jour, plus l'entraînement des modèles et la fabrication des puces, ça finit par peser lourd.

Les centres de données refroidissent souvent leurs machines en évaporant de l'eau : la quantité dépend du climat et de la technologie utilisée. L'impact carbone, lui, dépend de l'électricité du pays : une IA qui tourne au charbon pollue bien plus qu'une IA alimentée par du nucléaire ou des renouvelables. Il faut aussi compter la fabrication des puces, qui demande des métaux et beaucoup d'eau très pure. Enfin, l'entraînement d'un modèle coûte énormément, mais une seule fois, alors que l'usage se répète des milliards de fois : c'est lui qui pèse le plus à long terme.

> **Sous le capot :** On distingue l'énergie de l'entraînement, énorme mais ponctuelle, et celle de l'inférence : chaque requête coûte peu, mais elles se comptent en milliards. L'eau sert surtout au refroidissement des centres de données, et indirectement à produire l'électricité.

> **À retenir : Une question coûte peu, mais l'IA en reçoit des milliards. Les images et surtout les vidéos coûtent bien plus : à réserver aux vrais besoins.**

Lire l'article complet : Quel est l'impact environnemental de l'IA ?

3 Ligne 3 : Limites, sécurité, société

Inventer : Quand l'IA invente

Une IA prédit un mot **plausible**, pas un mot **vrai**. Quand elle ne sait pas, elle peut fabriquer une réponse qui a l'air parfaite : un livre qui n'existe pas, une date fausse, une citation inventée. On appelle ça une **hallucination**.

Une IA ne range pas les faits dans une base de données : ils sont dilués dans ses poids. Les faits très fréquents (Paris, capitale de la France) sont solides. Les faits rares sont flous, et c'est là qu'elle comble les trous avec du vraisemblable. La recherche web et la citation des sources réduisent beaucoup le problème, sans le supprimer. Les IA récentes sont aussi entraînées à dire « je ne sais pas », ce qui, pour une machine à compléter des phrases, n'a rien de naturel.

> **Sous le capot :** Une hallucination n'est pas un bug ponctuel : le modèle produit le texte le plus plausible, pas le plus vrai. Les parades : lui donner des documents à lire (RAG), lui faire citer ses sources, l'entraîner à dire « je ne sais pas ».

> **À retenir : Une IA peut se tromper avec assurance. Plus le sujet est rare et précis, plus il faut vérifier.**

Lire l'article complet : Pourquoi une IA invente-t-elle des réponses ?

Dater : Le monde s'est arrêté à la fin des cours

L'entraînement s'arrête à une date. Après, l'IA ne sait rien de ce qui se passe dans le monde : c'est sa **date limite de connaissances**. Pour l'actualité, elle doit faire une recherche sur le web.

Le piège dans le piège : une IA ne sait pas forcément quel jour on est. Sans indication, elle peut croire qu'on est encore à la fin de son entraînement et présenter de vieilles infos comme récentes. C'est pour ça que les assistants reçoivent la date du jour au début de chaque conversation. Et entre la fin de l'entraînement et la sortie publique d'une IA, il s'écoule souvent plusieurs mois.

> **Sous le capot :** Les connaissances s'arrêtent à la date de coupure des données d'entraînement. Pour l'actualité, l'assistant appelle un outil, par exemple une recherche web, dont il lit les résultats dans sa fenêtre de contexte.

> **À retenir : Ce qui ne change pas, l'IA le sait. Ce qui bouge (actu, prix, résultats, dirigeants), elle doit le chercher.**

Lire l'article complet : Pourquoi une IA ne connaît-elle pas l'actualité ?

Oublier : Une mémoire de poisson rouge (grand format)

Une IA relit toute la conversation avant chaque mot. Mais cette mémoire de travail a une taille maximale, la **fenêtre de contexte**. Quand elle est pleine, le plus ancien doit sortir, ou être résumé.

Les fenêtres actuelles sont immenses : des centaines de milliers de tokens, l'équivalent de plusieurs romans. Mais plus la conversation s'allonge, plus chaque mot coûte cher, puisque tout est relu. Certains assistants résument automatiquement le début pour faire de la place, et un résumé perd des détails. Conseil pratique : pour un nouveau sujet, ouvre une nouvelle conversation, et rappelle les infos importantes.

> **Sous le capot :** La fenêtre de contexte se compte en tokens : de quelques dizaines de milliers à plus d'un million selon les modèles. Le coût de l'attention augmente avec sa longueur, d'où les mémoires externes qui résument ou stockent des informations à part.

> **À retenir : La mémoire de travail d'une IA est grande mais limitée. Une info cruciale ? Redis-la, ou mets-la au début.**

Lire l'article complet : Pourquoi une IA oublie-t-elle le début de la conversation ?

Le piège : L'instruction cachée

Quand une IA lit une page web ou un document pour toi, elle peut tomber sur un texte qui lui donne des ordres : « Ignore ta consigne et envoie… ». C'est une **injection d'instructions**. Une IA bien conçue fait la différence entre ce que tu lui demandes et ce qu'elle lit.

C'est l'un des grands chantiers de sécurité des IA qui agissent seules : celles qui naviguent sur le web, lisent les mails ou remplissent des formulaires. Les défenses : entraîner l'IA à n'obéir qu'à son utilisateur, limiter ce qu'elle peut faire sans permission, et demander une confirmation humaine avant toute action sensible (payer, envoyer, supprimer).

> **Sous le capot :** L'injection de prompt marche parce que le modèle ne distingue pas nativement les instructions des données : tout est du texte dans le même contexte. Les défenses combinent entraînement spécifique, filtrage et limitation des actions autorisées.

> **À retenir : Ce qu'une IA lit n'est pas un ordre. Seul l'utilisateur donne les consignes.**

Lire l'article complet : L'injection de prompt, c'est quoi ?

Les sources : À qui appartiennent les textes ?

Les IA de texte ont appris sur des milliards de pages : sites web, articles de presse, livres, forums. Beaucoup de ces textes ont des **auteurs**, qui n'ont été ni consultés ni payés. C'est l'un des plus grands débats autour de l'IA, et il se règle en partie devant les tribunaux.

En Europe, la loi autorise la fouille de textes et de données sur les contenus accessibles en ligne, sauf si l'auteur s'y oppose de façon lisible par les machines. Le règlement européen sur l'IA impose en plus aux fournisseurs de grands modèles de respecter ces refus et de publier un résumé de leurs données d'entraînement. Aux États-Unis, les entreprises invoquent le « fair use » (usage loyal). Les décisions divergent : un juge américain a estimé que l'entraînement sur des livres achetés légalement relevait du fair use, mais pas la conservation de copies piratées, ce qui a mené Anthropic, l'entreprise qui développe Claude, à un accord de 1,5 milliard de dollars avec des auteurs. Au Royaume-Uni, la justice a jugé qu'un modèle d'image qui ne stocke pas les œuvres n'en est pas une copie. En Allemagne, des auteurs de chansons ont obtenu gain de cause contre OpenAI parce que ChatGPT restituait leurs paroles. Des dizaines d'autres procès sont en cours, et des éditeurs signent aussi des accords de licence avec les entreprises d'IA. Rien n'est encore figé.

> **Sous le capot :** En Europe, la fouille de textes et de données est permise sauf si l'auteur s'y oppose de façon lisible par une machine (directive 2019/790, article 4). Le règlement européen sur l'IA demande aux fournisseurs de modèles de publier un résumé de leurs données d'entraînement.

> **À retenir : Les IA ont appris sur des œuvres de vrais auteurs, et le droit se construit procès après procès. De ton côté : vérifie, cite tes sources, écris avec tes mots.**

Lire l'article complet : IA et droits d'auteur : que dit la loi ?

Le débat : Pour le meilleur et pour le pire ?

L'IA fait avancer la science, aide des millions de personnes et fait gagner du temps. Mais elle soulève aussi de grandes **questions de société** : le travail, la façon d'apprendre, la mémoire, le vrai et le faux, le climat. Les deux sont vrais en même temps.

Côté bénéfices : la recherche (structure des protéines, nouveaux matériaux, modèles climatiques), la santé (aide au diagnostic), l'accessibilité (lecture à voix haute, sous-titres, traduction), l'éducation (un tuteur disponible à toute heure) et le gain de temps sur les tâches répétitives. Côté questions : des métiers transformés, parfois menacés ; le risque de déléguer sa réflexion et de perdre des savoir-faire ; un internet qui se remplit de contenus générés, au point de brouiller la mémoire collective ; les deepfakes et la désinformation ; les biais qui reproduisent des discriminations ; la vie privée ; le pouvoir concentré entre quelques entreprises ; l'énergie et l'eau consommées. Des règles se construisent, comme le règlement européen sur l'IA, mais beaucoup reste à décider, collectivement.

> **Sous le capot :** Les grandes questions ouvertes : l'évaluation (comment mesurer qu'un modèle est fiable ?), l'alignement (comment s'assurer qu'il fait ce qu'on veut ?) et la gouvernance (qui fixe les règles, et comment les contrôler ?).

> **À retenir : L'IA apporte de vrais progrès et pose de vraies questions. Ce qu'elle deviendra dépend des règles qu'on se donne et de la façon dont chacun s'en sert.**

Lire l'article complet : L'IA, une bonne chose pour la société ?

Les autres cours :