



# COMMENT FONCTIONNE UN GRAND MODÈLE DE LANGUE ?

*Olivier Cappé, Claire Marc, Tout comprendre (ou presque)  
sur l'intelligence artificielle, 2025 © CNRS Éditions*

LES GRANDS MODÈLES DE LANGUE OU LLM (pour Large Language Model en anglais),  
NE "COMPRENENT" PAS LES TEXTES COMME LE FERAÏT UN HUMAIN. ILS LES  
REPRÉSENTENT SOUS FORME DE SÉQUENCES D'UNITÉS INDIVIDUELLES.

## QUELLES SONT CES UNITÉS ?

L'ENSEMBLE DES UNITÉS CHOISIES POUR LES LLM EST CONÇU POUR COUVRIR UNE VASTE  
GAMME DE TEXTES POSSIBLES. IL INCLUT :

au  
le est  
de une où

Les mots les plus courants  
d'une langue (conjonctions,  
articles, prépositions, etc.)

anti- -isme  
-able  
-tion pré-

Des morceaux de mots  
fréquemment utilisés  
(préfixes, suffixes, etc.)

& N t À  
v % 5

Tous les caractères  
(à l'exception de l'espace qui  
sert de séparateur)

UNE PHRASE EST ENSUITE DÉCOMPOSÉE EN UNE SÉQUENCE D'UNITÉS, EN PRIVILÉGIANT LES  
PLUS LONGUES CAR PLUS UNE UNITÉ EST LONGUE, PLUS ELLE A DE CHANCES DE CORRESPONDRE  
À UN MOT ENTIER OU À UNE PARTIE SIGNIFICATIVE DE CELUI-CI, CE QUI FACILITE  
LA COMPRÉHENSION DU CONTEXTE ET RÉDUIT L'AMBIGÜITÉ :

*"Les marguerites s'effeuillent au ralenti"*

EST UNE CHAÎNE DE CARACTÈRES QUI DEVIENT LA SÉQUENCE D'UNITÉ SUIVANTE :

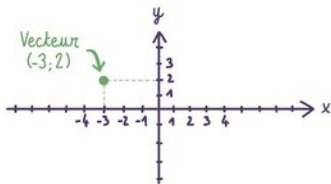
[ "Les", "mar#", "guer#", "ites", "s", "eff#", "eu#", "illent", "au", "ralenti", " ". ]

LE "#" DÉSIGNE UNE COUPURE QU'IL FAUDRA RECOLLER APRÈS TRAITEMENT PAR LE MODÈLE.

LE CARACTÈRE **MULTILINGUE** DES LLM NÉCESSITE DE CHOISIR UN **ENSEMBLE D'UNITÉS** AUSSI **"UNIVERSEL"** QUE POSSIBLE, AVEC DES GAMMES VARIANT ENTRE 30 000 ET 150 000 UNITÉS SELON LES MODÈLES.

## COMMENT REPRÉSENTER CES UNITÉS ET LEUR SENS ?

LA NOTION DE SENS EST DIFFICILE À FORMALISER EN INFORMATIQUE. UNE ALTERNATIVE CONSISTE À REPRÉSENTER CHAQUE UNITÉ PAR UN **VECTEUR** (un ensemble de valeurs numériques).



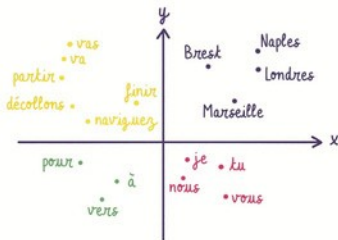
UN **VECTEUR** PEUT ÊTRE VISUALISÉ COMME UN **POINT** DANS UN ESPACE À **PLUSIEURS DIMENSIONS**.

Par exemple, dans un espace à 2 dimensions, un point est défini par une abscisse (x) et une ordonnée (y).

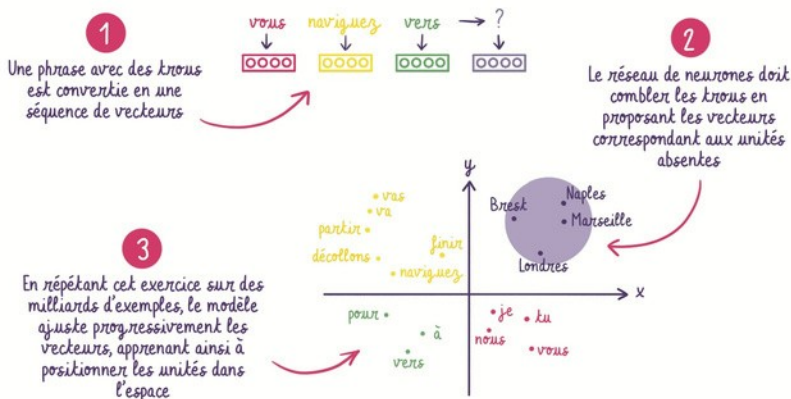
TOUTEFOIS, ÉTENDRE CETTE NOTION AUX MOTS EST PLUS DIFFICILE CAR ILS COMPTENT **BIEN PLUS QUE 2 DIMENSIONS** (500 à 1000).

CES DIMENSIONS ET LEURS INTERPRÉTATIONS SONT APPRISSES AUTOMATIQUEMENT EN S'APPUYANT SUR UN PRINCIPE LINGUISTIQUE CLÉ : DEUX UNITÉS SE RESSEMBLENT SI ELLES APPARAÎSSENT DANS DES CONTEXTES SIMILAIRES.

CETTE **RESSEMBLANCE** DEVIENT ALORS **MESURABLE** LORSQU'ON REPRÉSENTE LES UNITÉS SOUS FORME DE **VECTEURS**, CAR L'ON SAIT COMPARER DEUX VECTEURS ENTRE EUX (grâce à des outils mathématiques comme le **produit scalaire**) :



POUR APPRENDRE CES RESSEMBLANCES, ON JOUE AU "TEXTE À TROUS" AVEC UN RÉSEAU DE NEURONES ARTIFICIELS À PARTIR D'ÉNORMES QUANTITÉS DE TEXTES.



DANS CE CONTEXTE, LE MODÈLE APPREND QUE CERTAINS MOTS SE RESSEMBLENT CAR ILS PARTAGENT DES CONTEXTES SIMILAIRES :

| Pronoms                 | Verbes                                           | Prépositions            | Villes                                     | Ponctuation      |
|-------------------------|--------------------------------------------------|-------------------------|--------------------------------------------|------------------|
| TU<br>JE<br>ILS<br>VOUS | VAS PARTIR<br>M'ENVOLE<br>VONT FINIR<br>NAVIGUEZ | DE<br>POUR<br>À<br>VERS | MARSEILLE<br>LONDRES<br>MONTLUÇON<br>BREST | .<br>.<br>!<br>? |

OR, CHAQUE DIMENSION OUVRE LA POSSIBILITÉ DE QUANTIFIER UNE PROPRIÉTÉ DIFFÉRENTE, DE DISTINGUER OU DE RAPPROCHER DEUX UNITÉS (Londres et Marseille sont deux villes, mais distinctes). C'EST POURQUOI LES VECTEURS POSSÈDENT BIEN PLUS QUE 2 DIMENSIONS.

# MAIS COMMENT CAPTurer LE SENS D'UNE PHRASE ?

PRENONS L'EXEMPLE D'UN AVIS SUR UN FILM RÉDIGÉ EN UNE PHRASE :

*"Une musique terrible, mais la fin n'est vraiment pas folle."*

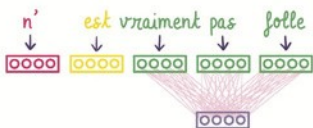
POUR ANALYSER AUTOMATIQUEMENT SI CET AVIS EST POSITIF OU NÉGATIF, IL FAUT D'ABORD DES DONNÉES D'EXEMPLES (avis classés comme positifs ou négatifs) PERMETTANT D'ENTRAÎNER UN RÉSEAU DE NEURONES, OÙ CHAQUE MOT EST TRANSFORMÉ EN UN VECTEUR NUMÉRIQUE.

SAUF QUE LE SENS DE CHAQUE MOT DÉPEND :

DU CONTEXTE



DES MOTS VOISINS

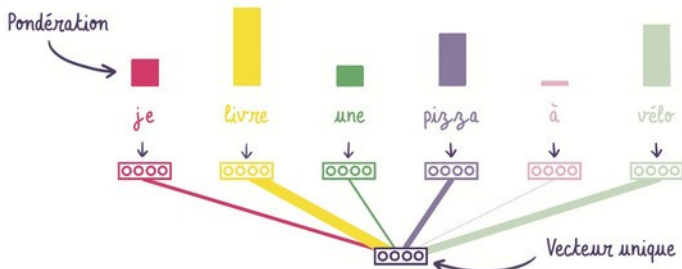


LES RÉSEAUX DE CONVOLUTION CAPTURENT CE CONTEXTE IMMÉDIAT. ILS ANALYSENT DES GROUPES DE MOTS PROCHES (appelés "fenêtres") ET CRÉENT DES VECTEURS ENRICHIS, PRENANT EN COMPTE LES INTERACTIONS LOCALES POUR AFFINER LA COMPRÉHENSION DE LA PHRASE.

## ET LES TRANSFORMEURS DANS TOUT ÇA ?

INTRODUITS EN 2018, LES TRANSFORMEURS AMÉLIORENT LA CONTEXTUALISATION EN PRENANT EN COMPTE NON PAS DES PETITES PARTIES D'UNE SÉQUENCE, MAIS L'ENSEMBLE DE CELLE-CI, GRÂCE À UN MÉCANISME CLÉ : L'ATTENTION.

L'ATTENTION AIDE LE MODÈLE À DÉTERMINER QUELS MOTS D'UNE PHRASE SONT LES PLUS SIGNIFICATIFS POUR COMPRENDRE LE CONTEXTE. CE MÉCANISME GÉNÈRE UN VECTEUR UNIQUE QUI POSSÈDE UNE PONDERATION SPÉCIFIQUE À CHAQUE MOT :



LES TRANSFORMEURS SONT ESSENTIELS DANS LES MODÈLES AVANCÉS DE TRAITEMENT DU LANGAGE, COMME GPT (le modèle de ChatGPT). D'AILLEURS, GPT VEUT DIRE :

- T pour "Transformeur" : la brique de base
- P pour "Pre-trained" (pré-entraîné) : l'entraînement sur des milliards de textes
- G pour "Generative" : le modèle complète les séquences avec cohérence

## LE MOT DE LA FIN

LES CONVOLUTIONS ET TRANSFORMEURS PERMETTENT DES MODÈLES TOUJOURS PLUS GRANDS, ATTEIGNANT DES CENTAINES DE MILLIARDS DE PARAMÈTRES, SOULEVANT PLUSIEURS QUESTIONS :



### Qualité des données

Quels biais et stéréotypes contiennent-elles ?  
Sont-elles libres de droit ?



### Coûts

Environnementaux (énergie)  
et financiers (plusieurs millions d'euros)



### Viabilité

Ces services, souvent gratuits, peuvent-ils être durables ?