

Mis à jour le 26/04/2024

S'inscrire

Formation LLM

3 jours (21 heures)

Présentation

Ces modèles linguistique basés sur des [transformateurs](#) sont des puissants outils qui vont répondre à une variété de tâches linguistiques. Notre formation LLM vous apprendra à utiliser ce modèle de langage.

Pendant ce cours, vous utiliserez diverses techniques, pratiques et playbooks pour créer des produits intégrant la puissance des modèles de langage.

Vous découvrirez en profondeur l'architecture Transformer et toutes ses variantes. De plus, vous développerez une intuition sur l'architecture et de son impact en fonction de vos décisions.

La formation vous enseignera à créer des [pipelines LLM](#) avancés pour regrouper des documents texte et explorer les sujets auxquels ils appartiennent.

Comme pour toutes nos formations, notre formation LLM vous sera présenté avec ses toutes dernières nouveautés (à la date de rédaction de l'article).

Objectifs

- Comprendre l'architecture des modèles Transformer sous-jacents tels que BERT et GPT
- Savoir comment sélectionner un modèle de langage pour vos besoins
- Comprendre comment fonctionnent les LLM
- Optimisez les LLM pour des applications spécifiques

Public visé

- Développeurs
- Professionnels IT

Pré-requis

- Connaissance de base en traitement automatique du langage naturel (TALN) et en modélisation linguistique
- Connaissance de base en informatique et en programmation
- Compréhension de la tokenisation

PROGRAMME DE NOTRE FORMATION LLM

Données d'entraînement

- Ingrédients d'un LLM
- Exigences en matière de données de pré-entraînement
- Ensembles de données de préformation populaires
- Prétraitement des données de formation
- Exploitation des caractéristiques des ensembles de données de préformation
- Problèmes de biais et d'équité dans les ensembles de données de préformation

Tokenisation, objectifs d'apprentissage et architectures

- Vocabulaire et tokenisation
- Objectifs d'apprentissage
- L'architecture
 - Architectures à codeur seul
 - Architectures codeur-décodeur
 - Architectures à décodeur seul
- Résumé

Adapter les LLM à votre cas d'utilisation

- Naviguer dans le paysage du LLM
 - Qui sont les fournisseurs de LLM ?
 - Les saveurs du modèle
- Comment choisir un LLM pour votre tâche ?
- Accès et chargement des LLM
- Stratégies de décodage

Génération de textes avec des modèles GPT

- Utilisation des modèles de génération de texte
- Introduction à l'ingénierie des messages-guides
- Ingénierie avancée des messages-guides
- Raisonner avec des modèles génératifs
- Vérification des résultats

Modèles multimodaux à langage étendu

- Transformateurs pour la vision
- Modèles d'intégration multimodaux
 - BLIP 1 : Connecter le texte et les images
- Rendre les modèles de génération de texte multimodaux
 - BLIP-2 : Comblent le fossé entre les modalités
 - Prétraitement des données multimodales
 - Cas d'utilisation 1 : sous-titrage d'images
 - Cas d'utilisation 2 : messages multimodaux basés sur le chat

Création de modèles d'intégration de texte

- Intégrer des modèles
- Qu'est-ce que l'apprentissage contrastif ?
- SBERT
- Création d'un modèle d'intégration
- Affiner un modèle d'intégration
- Apprentissage non supervisé

Génération améliorée par récupération (RAG)

- La nécessité de la RAG
- Scénarios typiques
- Le pipeline RAG
- RAG pour la gestion de la mémoire
- RAG pour la sélection d'exemples de formation en contexte
- RAG pour l'apprentissage du modèle
- Limites de RAG
- RAG vs. contexte long
- RAG vs fine-tuning

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format

numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.