Aligner les LLMs aux graphes de connaissances pour des explications adaptées.

Encadrant

Massinissa **HAMIDI**

Etudiants

MOKHTARI Dhia Elhak **LAKHMI** Hichem Billal **OUHENIA** Anis

Comment les LLMs peuvent-ils s'adapter selon les connaissances de chaque étudiant ?

By linking each student to a personalized Educational Knowledge Graphs (**EKGs**) where concepts are connected through **weighted prerequisite relationships**.

We can model not only what the student needs to learn, but also how strongly each concept depends on another using **weights**. Each concept node in the knowledge graph carries a **known** property indicating whether the student has mastered it.

This structure enables the system to identify the most critical knowledge gaps and prompt a Large Language Model **(LLM)** to generate tailored explanations that build only on the student's known concepts, while strategically avoiding or simplifying those not yet acquired.

Méthodologie.
Bridging Structured Domain Knowledge with Adaptive LLM Content Generation.

1. Student
Represents the learner. tracking his understanding via a **known** property on each concept node.

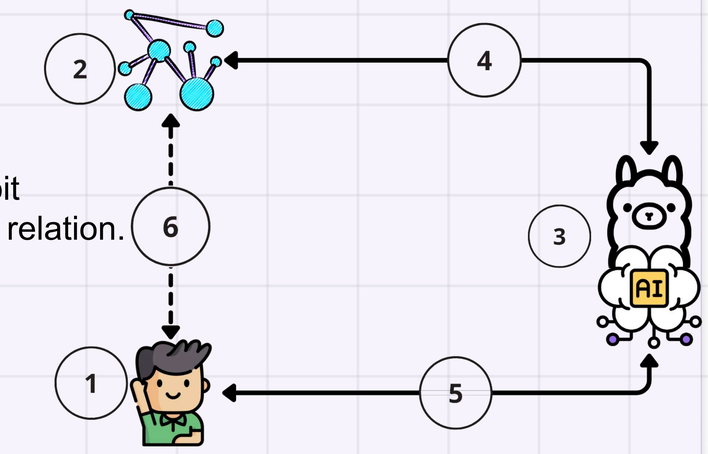
2. Knowledge Graph
Construit à l'aide de la méthodologie ACE, Chaque paire reçoit un score de relation de prérequis. avec des poids sur chaque relation.

3. LLM (Large Language Model)
LLM explains the Concepts handed.

4. Guidage via Graphe
Traverses the Graph. If missing prerequisites are found (known = 0), LLM is guided to explain first, in order of weight.

5. Interaction d'apprentissage
Cette réponse peut ensuite mettre à jour l'état (known) si l'étudiant valide l'apprentissage.

6. Feedback Loop
Les connaissances acquises par l'étudiant sont utilisées pour mettre à jour dynamiquement le graphe.



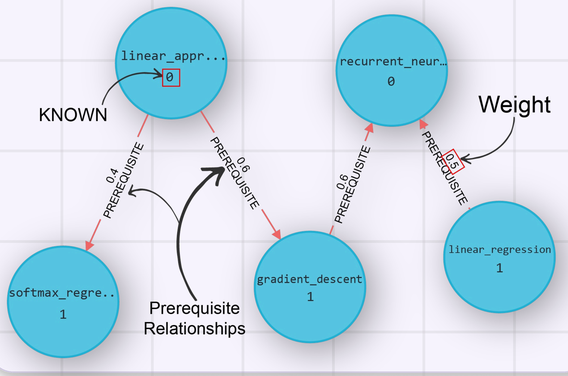
Learning Scenarios

Learning **Softmax Regression**

1. Alice selects Softmax Regression as her target concept.
2. The system checks her knowledge graph. It detects that Logistic Regression and Gradient Descent are not yet mastered.
3. The LLM first explains these missing prerequisites, then Softmax Regression.
4. Explanations and quizzes are personalized based on her current knowledge.
5. Alternatively, Alice could:
   📄 **Upload a PDF**, and the system would quiz/explain based on its content.
   ❓ **Ask about** any concept, Receive an explanation adapted to her level.

Concept Prerequisite Graph

- We use the **ACE** algorithm to build the graph:
- Concept pairs are extracted from course content using text-based analysis.
- Each pair is scored using:
  - **CSR** (Concept Semantic Relatedness)
  - **CER** (Concept Embedding Relatedness)
- Only high-scoring pairs are selected as potential **prerequisites**.
- Final edges are **expert-validated** to ensure correctness.
- The result is a **weighted**, acyclic, directed graph of prerequisite relations.



Large Language Model

Un LLM **(Large Language Model)** est un modèle de langage de grande taille entraîné sur de vastes corpus textuels.
Il est capable de comprendre et de générer du **langage naturel**, de répondre à des questions, d'expliquer des concepts ou encore de créer des quiz.
Dans ce projet, il est utilisé pour adapter dynamiquement les contenus pédagogiques au **niveau** de chaque **étudiant**.

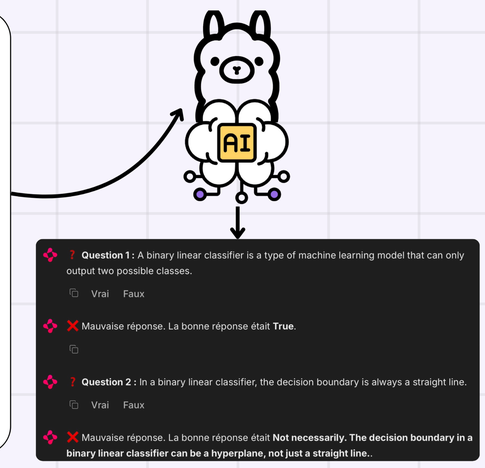
Few-Shot Prompting

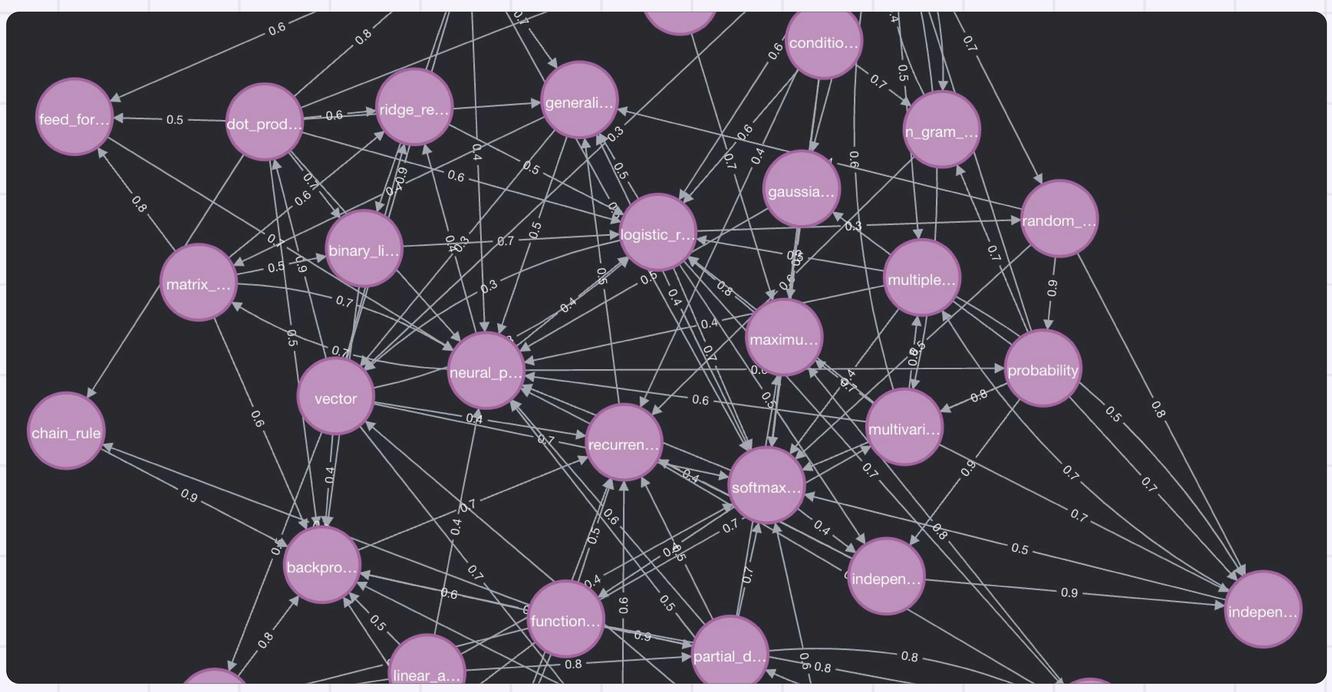
Nous guidons le **LLM** en lui fournissant quelques exemples de ce que nous attendons (ex. : format de quiz, style d'explication).

Cela permet de produire des contenus cohérents et adaptés à l'apprentissage

nous utilisons le few-shot prompting pour guider le modèle de langage (LLM)

```
quiz_prompt =(
f"Génère 3 questions Vrai/Faux avec les bonnes réponses, au format JSON comme ceci :\n"
f"[{{'question': '...', 'answer': 'Vrai'}}, ...] sur le concept '{concept}'.")
```



Résultats

	Knowledge Graph + LLM	ChatGPT (Modèle Générique)
Approche pédagogique	Guidée par des concepts structurés dans l'EKG (prérequis → compréhension progressive)	Réactive aux questions, sans parcours structuré ou suivi des prérequis
Génération de quiz/exemples	Générés dynamiquement à partir du concept + PDF + quiz	Peut en générer si demandé, mais sans lien avec un graphe/pdf
Suivi de progression	Oui, via known: 0/1 dans Neo4j, avec quiz + validation par score	Non : pas de structure de progression suivie nativement
Interaction	concept → quiz → explication → PDF → synthèse	on peut poser n'importe quelle question à tout moment
Objectif principal	Aider l'étudiant à maîtriser progressivement les concepts	Répondre à tout type de question, pas forcément dans un but d'apprentissage structuré
Confidentialité	Haute : tout peut tourner localement	Dépend des serveurs OpenAI (données peuvent transiter)