

Архитектура и механизмы адаптации агентов искусственного интеллекта для персонализации образовательного опыта

Науки о данных

Выполнил: Антонов Эдуард Кузьмич

Научный руководитель: Атаева Ольга Муратовна

к.т.н, старший научный сотрудник ФИЦ ИУ РАН

Проблема / исследовательский вопрос

Ключевые проблемы:

- Достоверность информации на уровне 75-80% ([MMLU](#))
- Отсутствие прозрачности логики принятия моделями решений.
- Ограниченное контекстное окно (32 - 128 тысяч токенов).

Исследовательский вопрос:

Как интеграция большой языковой модели с графом знаний в рамках агентной архитектуры позволит преодолеть ограничения монолитных моделей и создать диалоговую образовательную систему, поддерживающую адаптацию контента к уровню знаний, стилю обучения и образовательным целям пользователя?

Цель исследования

Разработать и экспериментально оценить агентную архитектуру, объединяющую большие языковые модели и граф знаний для создания адаптивной образовательной диалоговой системы.

Объект исследования:

Образовательный процесс с использованием генеративного искусственного интеллекта и графовых структур данных.

Предмет исследования:

Архитектура агентов на базе больших языковых моделей.

Обучающие машины

- Развитие персонализированного обучения.
- Ранние адаптивные обучающие системы.
- Модели отслеживания знаний.
- Большие языковые модели в образовании.
- Интеграция языковых моделей с графами знаний
- Агенты искусственного интеллекта.



AutoTutor Mark II
Western Design, 1960

Новизна

Агентная архитектура, реализующая динамическое взаимодействие языковой модели с графом знаний и долговременной памятью. Такой подход позволяет преодолеть ограничения статических методов извлечения информации, повысить релевантность генерируемого материала, обеспечить адаптивность диалога и персонализацию образовательных траекторий в сложных доменных областях.

Актуальность

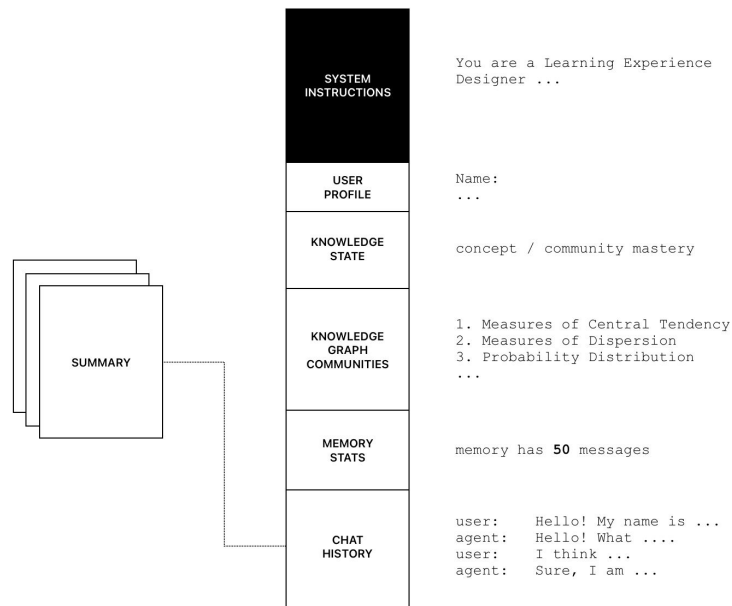
Активная цифровизация образования и интеграция генеративного ИИ создают предпосылки для разработки инновационных архитектурных решений, позволяющих создавать адаптивные образовательные системы, способные эффективно работать в условиях существующих ограничений языковых моделей.

Гипотеза исследования

Разработка агентной архитектуры, объединяющей большие языковые модели с графом знаний, обеспечит создание образовательной диалоговой системы, способной адаптироваться к индивидуальному уровню знаний и стилю обучения, повышать точность предоставляемой информации по сравнению со стандартными методами генерации с дополненной выборкой, а также поддерживать долгосрочное взаимодействие посредством эффективного управления контекстным окном модели.

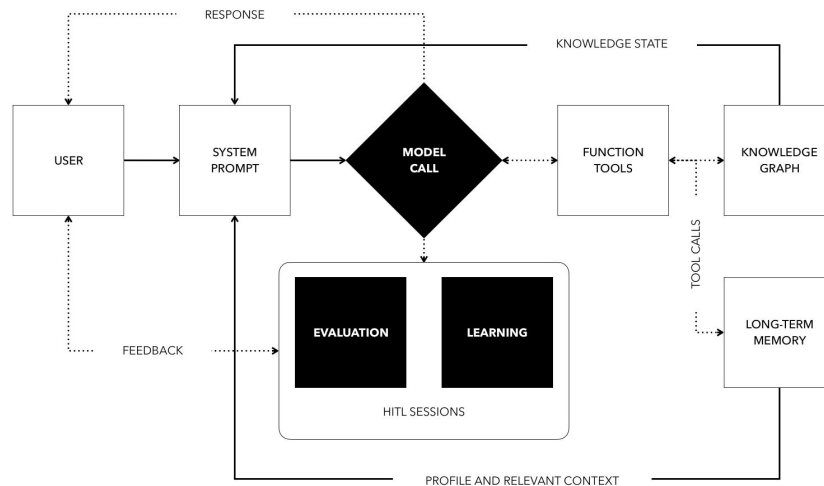
Задачи исследования

1. Разработать эффективные методы генерации с дополненной выборкой.
 - оптимизация запросов;
 - итеративное уточнение;
 - динамический системный промпт.
2. Формализовать предметную область в виде графа знаний.
 - иерархии и предшествования.
3. Реализовать прототип и провести оценку.
 - извлечение информации;
 - экспертная аннотация диалогов.



Методология исследования

- Jupyter Notebooks.
- Сравнение векторного и графового RAG (Top-1 Accuracy, Precision@K).
- Реализация интерфейса (FastAPI, React).
- Внедрение логирования внутренних процессов.
- Пилотное тестирование системы.
- Экспертная аннотация.



Источники данных

1. [Machine Learning Glossary](#) (Google) для создания графа.
2. Коллекция вопросов для оценки извлечения информации.
3. История взаимодействия с системой.

	term	content
0	ablation	A technique for evaluating ...
1	A/B testing	A statistical way of ...
2	accelerator chip	A category of specialized ...
3	accuracy	The number of correct ...
4	action	In reinforcement learning, ...
5	activation function	A function that enables ...
6	active learning	A training approach ...
7	agent	In reinforcement learning, ...
...

	category	question	target
0	direct	ReLU	relu
1	direct	Activation Function	activation function
..	
99	single	What is supervised learning?	supervised learning
100	single	How do you detect overfitting in a model?	overfitting
...	
199	multi	Can you explain how backpropagation utilizes gradient ...	backpropagation, ...
200	multi	What roles does ReLU play as an activation function in ...	relu, activation ...
...	

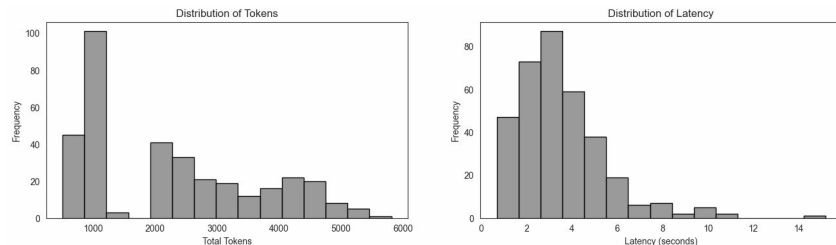
Дизайн эксперимента

Точность извлечения информации:

- One-shot (один запрос)
- Iterative (итеративное обращение)

Оценка качества:

Экспертная аннотация логов системы по критериям адаптивности, связности, использования контекста и корректного вызова инструментов.



Retriever	Top-1 Accuracy (%)	Precision@3 (%)
graph-based	83.3 \pm 3.8	74.6 \pm 1.6
multilingual-e5-base	67.7 \pm 0.72	79.9 \pm 4.2
text-embedding-3-large	69.7 \pm 1.9	82.1 \pm 1.3

Criteria	Rater A	Rater B	Cohen's Kappa
personalisation	2.8	2.7	0.57
coherence	2.9	2.8	0.78
relevance	2.9	2.8	0.78
tool use	2.4	2.4	0.66
context use	2.7	2.5	0.59

Результаты исследования

Точность извлечения информации:

- Графовый метод достиг Top-1 Accuracy $\approx 83,3\%$ (94,5 % итеративно) с минимальными рекурсивными ошибками (3,5 %)
- Векторные методы показали точность 68–72 % с более высоким Precision@K3 и риском рекурсивных ошибок (22,5–29 %)

Анализ диалогового взаимодействия:

- Система способна интегрировать ключевые детали из долговременной памяти.
- Образовательный контент динамично адаптируется под индивидуальные особенности обучающегося (цели, предпочтения, текущий уровень знаний)
- Необоснованная оценка ответов пользователя.
- Обнаружены отдельные проблемы с обновлением статуса знаний и интерпретацией терминов.

Практическая значимость работы

- Основа для полнофункциональных систем.
- Open Source Model Context Protocol (MCP).

Демо стенд 1:

Sreamlit

- базовый агент.

Демо стенд 2:

React + FastAPI

- агент с планированием.

GitHub Репозиторий

*В процессе доработки

(Добавить ссылки)

learning session

Introduction to Machine Learning

Machine Learning (ML) is a pivotal technology in modern data science. As aspiring data scientists, understanding ML will serve as a cornerstone in your career preparations, especially with interests in technical aspects. ML enables computers to learn from data and make predictions or decisions without being explicitly programmed for specific tasks. This can range from predicting weather patterns to personalizing online content.

Historical Context

The journey of machine learning began in the mid-20th century. Alan Turing, a pioneer in computing, posed the question: "Can machines think?" This question laid the groundwork for ML's evolution. In 1959, Arthur Samuel coined the term "machine learning" during his work on game playing software for checkers. Developments continued with significant advancements, such as:

- **1940s:** The "nearest neighbor" algorithm, which illustrated the concept of pattern recognition.
- **1980s:** Introduction of neural networks, which mimic human brain processing, leading to significant breakthroughs across various domains.
- **1997:** IBM's Deep Blue defeated world chess champion Gary Kasparov, showcasing the potential of ML.
- **2000s and beyond:** Advancements in computational power and data availability have exponentially grown ML's applications.

Core Concept

At its heart, machine learning involves training models to identify patterns within data, making predictions or decisions without direct programming for specific outcomes. This process requires several key steps, including:

1. **Data Collection:** Gathering data from which the model can learn.
2. **Feature Extraction:** Identifying the most relevant characteristics or features within the data.
3. **Model Training:** Teaching the model via algorithms to perform certain tasks like classification or regression.
4. **Validation:** Ensuring model accuracy by testing it on separate sets of data.

Significance in the Modern World

Machine learning propels various fields, revolutionizing how industries operate. For example, in healthcare, ML models analyze medical data for diagnostic purposes, improving the efficiency of diagnosis and treatment plans. In the field of finance, algorithms assess real-time data to predict market trends, enhancing decision-making processes.

This foundational understanding will segue into exploring the specific types and applications of machine learning. Next, we will delve into key components such as supervised, unsupervised, and reinforcement learning, each of which supports different tasks and scenarios. Real-world applications will further illustrate these concepts, linking them with practical examples to solidify your comprehension of the topic.

Continue

Profile:

Name: Alan

Goals: Becoming a Data Scientist

Interests: Learning the fundamentals of Machine Learning and preparing to apply for a bachelor's degree

Preferences: Focus on technical aspects with simple examples and detailed math explanations

Current Learning Session:

1. Introduction to Machine Learning

Introduce learners to the basic concept and historical evolution of machine learning as both a practical tool and a field of study.

2. Main Body - Core Concepts and Examples of Machine Learning

Enable learners to understand the fundamental processes, types, and applications of machine learning, as well as recognize the broader implications of its use.

Апробация исследования

IEEE 6th International Conference on Robotics, Intelligent Control and Artificial Intelligence

China, December 6–8, 2024.

Доклад представлен, работа принята для публикации, ожидается размещение в сборнике материалов конференции. (SCOPUS)

EUDSCILEARN 2024 Международная научно-практическая конференция «Трансформация образования: современные технологии, инновационные подходы и лучшие практики»

Россия, 24–25 октября 2024 г.

Доклад представлен, работа принята для публикации, ожидается размещение в сборнике материалов конференции. (РИНЦ)

СПИСОК КЛЮЧЕВЫХ ИСТОЧНИКОВ

1. Watters A. Teaching Machines: The History of Personalized Learning. – Cambridge, MA: MIT Press, 2021. – DOI: 10.7551/mitpress/12262.001.0001.
2. Doignon J.-P., Falmagne J.-C. Knowledge Spaces and Learning Spaces [Электронный ресурс] // arXiv preprint arXiv:1511.06757. – 2015. – URL: <https://doi.org/10.4850/arXiv.1511.06757> (Дата обращения: 01.12.2024)
3. Abu-Rasheed H., Weber C., Fathi M. Knowledge Graphs as Context Sources for LLM-Based Explanations of Learning Recommendations [Электронный ресурс] // arXiv preprint arXiv:2403.03008. – 2024. – URL: <https://arxiv.org/abs/2403.03008> (Дата обращения: 01.12.2024)
4. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Kuttler H., Lewis M., Yih W., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks [Электронный ресурс] // arXiv preprint arXiv:2005.11401. – 2020. – URL: <https://doi.org/10.48550/arXiv.2005.11401> (Дата обращения: 01.12.2024)
5. Yao S. et al. ReAct: Synergizing Reasoning and Acting in Language Models [Электронный ресурс] // arXiv preprint arXiv:2202.03629. – 2022. – URL: <https://doi.org/10.48550/arXiv.2210.03629> (Дата обращения: 01.12.2024)
6. Packer C., Wooders S., Lin K., Fang V., Patil S.G., Stoica I., Gonzalez J.E. MemGPT: Towards LLMs as Operating Systems [Электронный ресурс] // arXiv:2310.08560v2. – 2024. – URL: <https://doi.org/10.48550/arXiv.2310.08560> (Дата обращения: 01.12.2024)

**Спасибо
за внимание!**

