M04 - GenAl Framewoks

Stand: 03.2025

1 Übersicht GenAl Frameworks

Aspekt	LangChain	Llama Stack	Haystack	LlamaIndex
Hauptfokus	Vielseitiges Framework für verschiedene NLP-A nwendungen	Standardisierung von Kernbausteinen für generative KI- Anwendungen	Entwicklung skalierbarer LLM-Anwendungen mit Fokus auf Suche und RAG	Datenframework für LLM- Anwendungen mit Fokus auf Indexierung und Abruf
Architektur	Modular mit Ketten, Agenten, Prompts und Speicher	APIs für Inferenz, Sicherheit, Speicher, Agenten und Bewertung	Pipeline-basiert mit Komponenten für Dokumentenverarbeitung, Retrieval und Generierung	Modulare Struktur mit Fokus auf Datenindexierung und - abfrage
Stärken	Flexibilität, Anpassbarkeit, breites Anwendungsspektrum	Einheitliche APIs, nahtloser Übergang zwischen Entwicklung und Produktion	Einfachheit, Skalierbarkeit, gute Dokumentation	Umfangreiche Datenverarbeitung, fortschrittliche Vektorspeicherfunktionen
Datenverarbeitung	Unterstützt verschiedene Datenquellen und Multimodalität	Fokus auf Llama- Modellfamilie, einschließlich spezialisierter Modelle	Effiziente Verarbeitung großer Datensätze, starke Retrieval-Fähigkeiten	Umfassende Datenkonnektor- Unterstützung, fortschrittliche Indexierungsfunktionen

Aspekt	LangChain	Llama Stack	Haystack	LlamaIndex
Einsatzbereich	Allgemeine NLP- Aufgaben, komplexe Workflows	Spezifisch für Llama- basierte Anwendungen	Suchsysteme, RAG- Anwendungen, Produktionsumgebungen	Kontextbasierte KI- Anwendungen, RAG- Implementierungen
Entwickler- freundlichkeit	Erfordert tieferes Verständnis für komplexe Anwendungen	Zielt auf einfache Integration und einheitliche Entwicklererfahrung ab	Einfach zu verstehen und zu erweitern, gute Dokumentation	Einfach zu bedienen, aber weniger umfangreich als LangChain
Skalierbarkeit	Gut für verschiedene Projektgrößen geeignet	Designed für nahtlosen Übergang von lokaler Entwicklung zu Cloud- Deployments	Optimiert für große Datensätze und Produktionsanwendungen	Gut skalierbar, besonders für Datenindexierung und - abfrage
Community und Support	Große, aktive Community mit umfangreicher Dokumentation	Wachsende Community, fokussiert auf Llama- Ökosystem	Starke Community- Unterstützung, gute Dokumentation	Wachsende Community, gute Dokumentation
Leistung	Effizientes Ressourcenmanagement durch optimierte Workflows	Speziell für Llama- Modelle optimiert	Leichtgewichtig und effizient, besonders bei RAG-Anwendungen	Effizient bei Datenindexierung und -abfrage
Anpassbarkeit	Hochgradig anpassbar, aber komplex für Anfänger	Fokussiert auf Standardisierung innerhalb des Llama- Ökosystems	Gute Balance zwischen Anpassbarkeit und Einfachheit	Flexibel, besonders bei Datenverarbeitungsaspekten

2 Gründe für LangChain

Für einen die Anwendung von GenAl ist **LangChain** zu empfehlen. Gründe für diese Empfehlung sind:

- **Breite Modellunterstützung**: LangChain unterstützt eine Vielzahl von großen Sprachmodellen (LLMs) wie OpenAl, Hugging Face, Cohere und mehr. Dies ermöglicht es Einsteigern, mit verschiedenen Modellen zu experimentieren und zu verstehen, wie sie funktionieren.
- Flexibilität und Vielseitigkeit: LangChain ist sehr flexibel und kann für eine Vielzahl von Anwendungsfällen eingesetzt werden, z. B. Chatbots, Frage-Antwort-Systeme, Agents und Dokumentenanalyse. Dies gibt Einsteigern die Möglichkeit, unterschiedliche GenAl-Anwendungen kennenzulernen.
- **Umfangreiche Dokumentation und Community**: LangChain hat eine große und aktive Community sowie eine umfangreiche Dokumentation. Dies ist besonders für Einsteiger hilfreich, da sie leicht Ressourcen, Tutorials und Unterstützung finden können.
- **Einfache Integration**: LangChain ermöglicht die einfache Integration mit externen Tools, APIs und Datenquellen. Dies ist nützlich, um Einsteigern zu zeigen, wie sie reale Daten und Dienste in ihre Anwendungen einbinden können.
- **Gute Balance zwischen Einfachheit und Funktionalität**: Während LangChain mächtig und flexibel ist, bietet es auch eine relativ einfache API, die für Einsteiger zugänglich ist. Es gibt viele Beispiele und Vorlagen, die den Einstieg erleichtern.
- **Zukunftssicherheit**: Da LangChain ein allgemeines Framework ist, das nicht auf ein bestimmtes Modell beschränkt ist, können Einsteiger die erlernten Konzepte auf zukünftige Modelle und Technologien anwenden.

3 Vergleich mit anderen Frameworks

- Llama Stack: Zu spezialisiert auf Llama-Modelle, was für Einsteiger einschränkend sein könnte.
- Rivet: Visuelles Tool, das zwar benutzerfreundlich ist, aber weniger allgemeine Konzepte vermittelt.
- Haystack: Fokussiert auf NLP und Frage-Antwort-Systeme, was für Einsteiger zu eng sein könnte.

Fazit

LangChain ist die beste Wahl für einen GenAl-Kurs, da es eine breite Palette von Anwendungsfällen abdeckt, flexibel ist und eine große Community sowie umfangreiche Dokumentation bietet. Es ermöglicht Einsteigern und Fortgeschrittenen, die Grundlagen der Generativen KI zu erlernen und gleichzeitig praktische Erfahrungen mit verschiedenen Modellen und Tools zu sammeln.