

Advanced Generative AI: Building LLM Applications



Benchmark and Evaluation of LLM Capabilities: Part 1



Quick Recap



- What is the primary purpose of fine-tuning Language Models (LLMs), and how is it achieved?
- Why is hyperparameter adjustment crucial in the fine-tuning process of LLMs, and when are these hyperparameters set?

Engage and Think



Imagine being part of a team that is responsible for analyzing a massive amount of user-generated content, reviews, and feedback from various markets worldwide. To make informed decisions, you need to summarize customer sentiments, address language barriers through efficient translation, and utilize chatbots to engage with users in real time.

What are the challenges you might face?

Learning Objectives

By the end of this lesson, you will be able to:

- 🔗 Apply the approach used in text summarization for condensing lengthy content
- 🔗 Utilize LLM models for translating content
- 🔗 Build a chatbot or virtual assistant using LLMs for automating tasks or retrieving information
- 🔗 Analyze LLM applications for understanding sentiments in text data





LLM Capabilities

LLM Capabilities

The four major capabilities of LLMs are:



Summarization

Content
translation

Chatbots and
virtual assistants

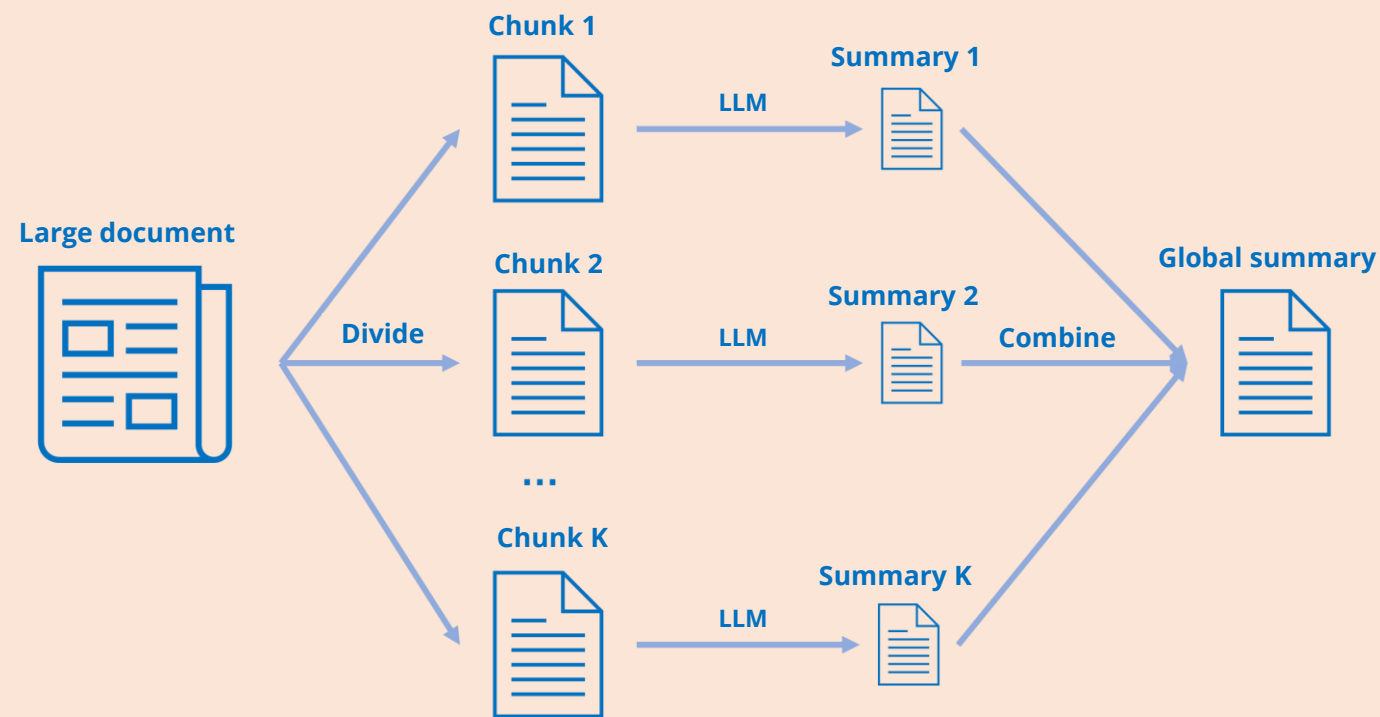
Sentiment
analysis



Summarization

Summarization

LLM summarization is a technique that uses LLMs to generate summaries of long texts or documents.



- The **refine** method breaks down a lengthy text into smaller chunks that fit the model's context, summarizes each chunk, concatenates the summaries, and uses the LLM to create a final summary.
- Another method combines individual document summaries from the **Map** step into a cohesive text string.

Summarization

LLM summarization can be used to:

Control tone and writing style

Obtain semantically-coherent topics

Generate distinct topic titles

Benefits of LLM Summarization

Efficiency

LLMs excel at swiftly processing and summarizing vast amounts of information, outperforming traditional manual methods.

Objectivity

LLM-based summarization provides a dynamic and objective approach, minimizing the influence of human perspectives and cognitive limitations.

Scalability

LLMs handle large volumes of data and easily scale for summarizing long documents or multiple sources of information.

Benefits of LLM Summarization

Accuracy

Achieving consistently increasing accuracy, LLMs are trained on massive datasets to deliver precise summaries.

Low-latency responses

Leveraging modern technology, LLMs generate rapid, low-latency responses, meeting real-time summarization needs.

Versatility

Applying their capabilities across various domains such as business, legal, and academia, LLMs offer a versatile solution for summarizing different types of content.

Limitations of Summarization

Quality of summaries

LLMs may not consistently generate accurate or unique insights, which can be crucial in some applications.

Complexity of datasets

LLMs can face challenges in training and deploying applications, especially when working with large and complex datasets.

Training costs

LLMs may incur prohibitive training costs, posing a challenge for smaller organizations or those with limited resources looking to adopt the technology.

Response latency

LLMs may exhibit high response latency, a concern for real-time applications such as summarizing live events or providing customer support.

Limitations of Summarization

Context limit

LLMs may encounter limitations in understanding and retaining context during summarization, resulting in a loss of information in the final summary.

Prompt fragility

LLMs can be sensitive to the choice of prompts, presenting challenges in developing reliable and consistent summarization solutions.

Repeatability

LLM summarization results can be challenging due to the inherent stochastic nature of the models and the choice of hyperparameters.

Summarization: Industrial Applications

Finance: LLMs summarize market reports and data.

Law: LLMs summarize legal documents and precedents.

Media and entertainment: LLMs summarize articles.

Marketing and advertising: LLMs summarize and generate marketing content.

Healthcare and enterprises: LLMs provide executive-level summaries of various functions.

Demo: Text Summarizer



Duration: 25 minutes

Overview:

In this demo, you will create an Arxiv paper summarizer that downloads a paper as a PDF, reads it, and generates a summary. You can use **The Impact of Generative Artificial Intelligence** paper as an example for demonstration.

Note

Please download the solution document from the Reference Material Section and follow the Jupyter Notebook for step-by-step execution.

DEMONSTRATION

Quick Check



How does the **refine** method work in LLM summarization?

- A. Break-Summarize-Concatenate
- B. Combine-String-Summarize
- C. Segment-Refine-Map
- D. Summarize-Combine-Refine



Content Translation

Content Translation

LLMs are increasingly used for content translation, offering several advantages and capabilities.



Content Translation

LLMs possess improved context awareness, allowing for a better understanding of:



Idiomatic
expressions

Cultural
nuances

Regional
variations

Benefits of Content Translation

Multilingual support

LLMs can work with multiple languages to foster global communication and information access.

Improved context awareness

LLMs have a strong ability to understand and generate text based on context, resulting in more accurate and culturally appropriate translations.

Reduced bias

LLMs provide an opportunity to address biases that may exist in the model's default settings, which leads to more fair and unbiased translations.

Efficiency

LLMs offer efficient and cost-effective language translation services to reduce the need for manual intervention and speed up processes.

Benefits of Content Translation

Consistency and quality

LLMs offer consistency and quality in various tasks, including content creation, quality assurance, and data analysis.

Accessibility and inclusivity

LLMs can translate content into multiple languages, making information accessible to a global audience.

Personalized assistance

LLMs have the potential to offer personalized assistance and recommendations to enhance the user experience.

Limitations of Content Translation

Accuracy and quality

LLMs are not perfect and can occasionally produce incorrect translations.

Biases

LLMs may have biases in their default settings, which can lead to unintentional biases in translations.

Legal and regulatory compliance

LLM-based language translation presents challenges when working with sensitive or regulated content.

Context awareness

LLMs may struggle with understanding and retaining context during translation.

Content Translation: Industrial Applications

Global businesses: LLMs are transforming the way documents are translated for global businesses and overcoming language barriers.

E-commerce: LLMs are highly valuable for e-commerce platforms in translating product descriptions and customer reviews automatically.

Marketing and advertising: LLMs create multilingual content, such as dynamic user notifications, localized marketing materials, or personalized messages.

Travel and tourism: LLMs provide real-time, high-quality translations that are often on par with those produced by human translators.

Language service industry: LLMs improve the efficiency and quality of machine translation, reducing the need for manual intervention and speeding up processes.

Quick Check

Which of the following capabilities of Large Language Models (LLMs) is enhanced due to their improved context awareness in content translation?

- A. Lexical disambiguation
- B. Syntactic parsing
- C. Idiomatic expressions
- D. Semantic role labelling

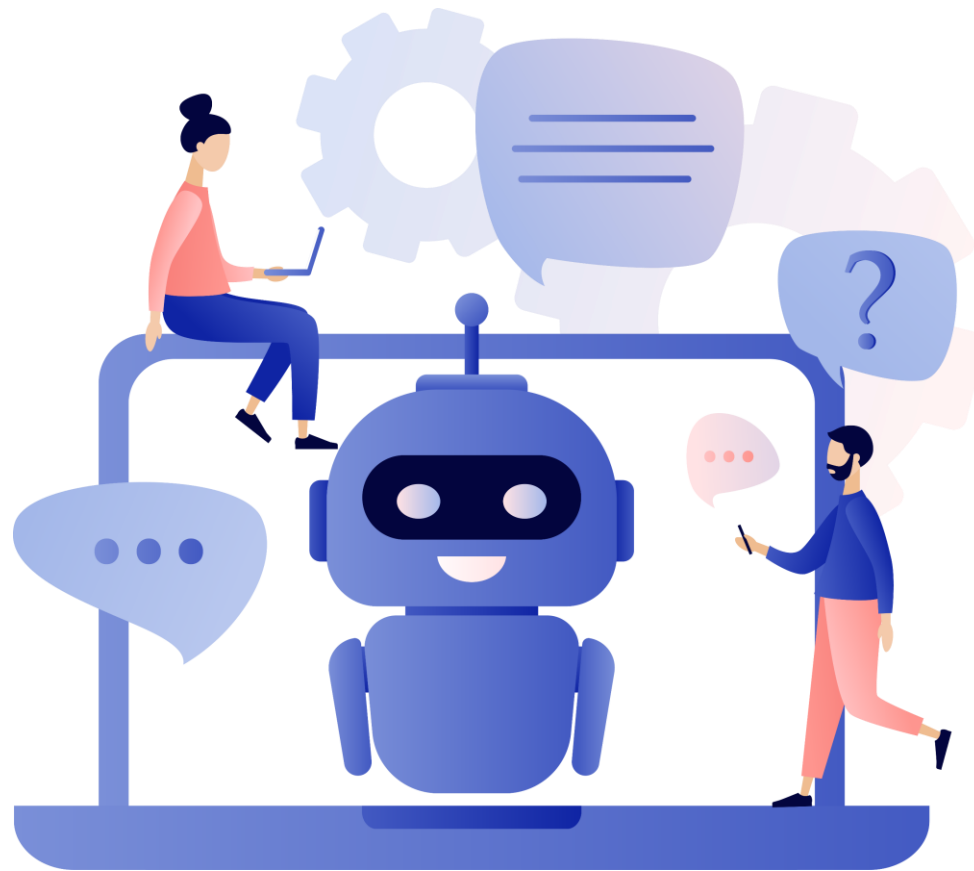




Chatbots and Virtual Assistants

Chatbots and Virtual Assistants

LLMs have enhanced conversational AI systems, allowing chatbots and virtual assistants to engage in more natural, context-aware, and meaningful conversations with users.



Benefits of Chatbots and Virtual Assistants

Cost savings

LLM-powered chatbots and virtual assistants can handle simple service requests to reduce the need for humans and lower costs.

Improved user engagement

LLM enables chatbots and virtual assistants to provide personalized, interactive experiences that engage users and keep them coming back.

Increased accessibility

LLM can make chatbots and virtual assistants more accessible to people with disabilities, such as those who are visually impaired or have limited mobility.

Real-time assistance

LLM-powered chatbots and virtual assistants can provide real-time assistance to users to help them resolve issues quickly and efficiently.

Benefits of Chatbots and Virtual Assistants

Multilingual support

LLM enables chatbots and virtual assistants to understand and respond to requests from global customers in multiple languages.

Enhanced customer support

LLM-powered chatbots are handling routine inquiries, troubleshooting issues, and providing relevant information with remarkable accuracy.

Personalized assistance

LLM can make chatbots and virtual assistants more accessible to people with disabilities, such as those who are visually impaired or have limited mobility.

24/7 availability

LLM-powered chatbots can operate round the clock, providing support and information at any time to improve customer service and satisfaction.

Chatbots and Virtual Assistants: Industrial Applications

Customer support: LLMs enhance customer service with 24/7 support, personalized recommendations, and reduced response times.

E-commerce: LLM-powered chatbots streamline purchases, improve efficiency, and enhance the overall customer experience.

Healthcare: LLM-powered virtual assistants automate tasks, handle inquiries, and offer real-time support.

Retail: LLM-powered chatbots are used for natural language conversations, understanding queries, and improving customer support.

Finance: LLM-powered chatbots are used for personalized trading, efficient onboarding, market predictions, and smart wealth management.

Demo: MultiPDF QA Retriever with FAISS and LangChain



Duration: 30 minutes

Overview:

In this demo, you will learn how to use LangChain to create a MultiPDF retriever with FAISS. This demo is performed on new generative AI research paper PDFs. You will understand how to load and process documents, create a database, make a retriever, create a chain, and use the retriever to ask questions and get answers.

Note

Please download the solution document from the Reference Material Section and follow the Jupyter Notebook for step-by-step execution.

Quick Check



Which of the following benefits are facilitated by Large Language Models (LLMs) in conversational AI systems?

- A. Semantic mapping
- B. Lexical disambiguation
- C. Syntactic parsing
- D. Personalized assistance



Sentiment Analysis

Sentiment Analysis

LLMs can understand the context in which words and phrases are used and perform sentiment analysis in real time.



Benefits of Sentiment Analysis

Improved decision-making

Companies can process vast amounts of data, enabling them to track brand sentiment and make data-driven decisions.

Personalized services

Companies can enhance their understanding of customer sentiment and personalize their services accordingly.

Improved customer support

Companies can respond promptly to customer needs, identify potential pain points, and provide more empathetic support.

Limitations of Sentiment Analysis

Cultural nuances

LLMs may struggle to capture cultural and contextual nuances that influence sentiment analysis, leading to incorrect sentiment predictions.

Sentiment ambiguity

LLMs may struggle with the inherent ambiguity in human language, leading to the potential misinterpretation of typos, synonyms, abbreviations, and mixed sentiments.

Lack of individual variation

LLMs can struggle with capturing individual variations in sentiment expression, leading to less accurate sentiment analysis.

Demo: Sentiment Analysis



Duration: 30 minutes

Overview:

In this demo, we aim to utilize two different tools to analyze the sentiment of a given text.
The two tools are:

- OpenAI's GPT-3 API
- TextBlob

Note

Please download the solution document from the Reference Material Section and follow the Jupyter Notebook for step-by-step execution.

Guided Practice



Overview

Duration: 15 minutes

In this activity, you will engage in a hands-on demonstration using LangChain and FAISS to create a MultiPDF retriever. By following the provided instructions, you will gain practical experience with these tools, enhancing your proficiency in leveraging LangChain and FAISS.

GUIDED PRACTICE

Key Takeaways

- LLM summarization can control tone and writing style, obtain semantically coherent topics, and generate distinct topic titles.
- LLMs are increasingly used for content translation, offering several advantages and capabilities.
- LLMs have enhanced conversational AI systems, allowing chatbots and virtual assistants to engage in more natural and meaningful conversations with users.
- LLMs can understand the context in which words and phrases are used and perform sentiment analysis in real time.



Q&A

