

Large Language Models



gollnickdata.de

Large Language Models

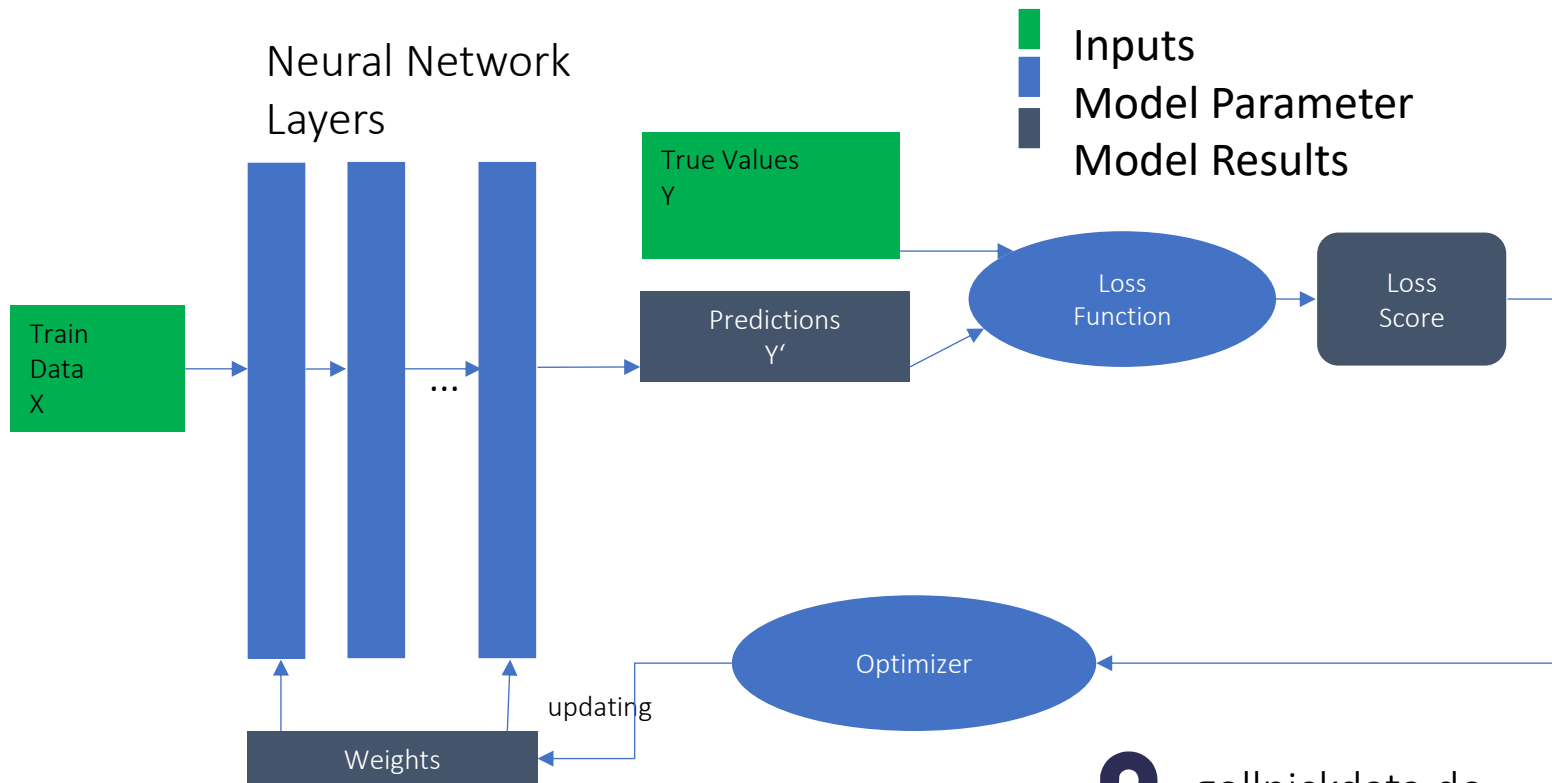
Introduction

- Type of artificial intelligence model
- Designed to understand, generate, and manipulate natural language text
- Trained on large (text) datasets
- Can perform various language tasks like translation, summarization, text generation, ...
- Capabilities improved dramatically in the last years
- Based on Deep Learning, specifically Transformers



Large Language Models

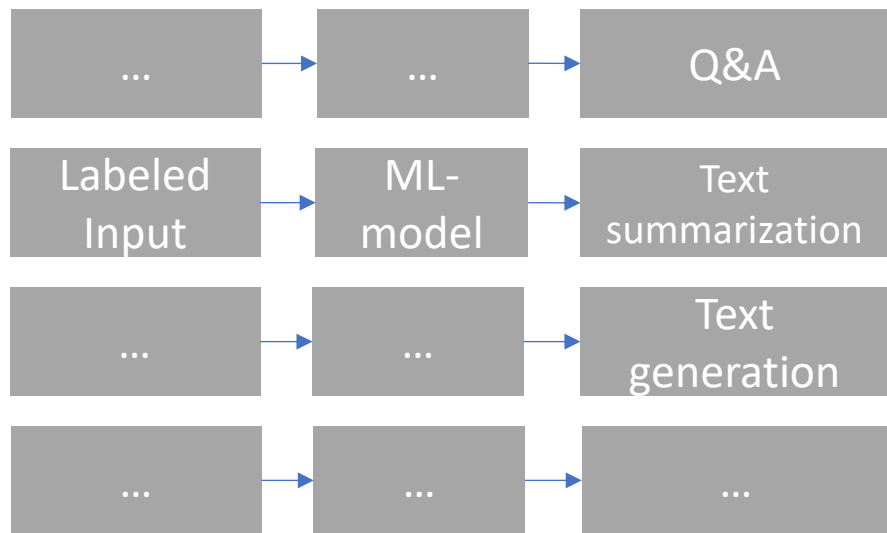
Deep Learning



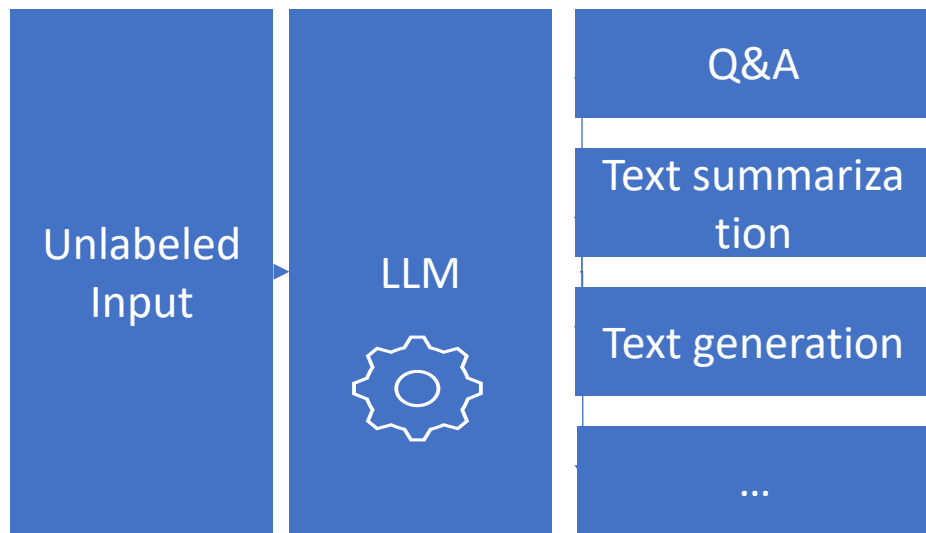
gollnickdata.de

Large Language Models

Difference to Classical Models



Classical ML-models



Large Language Model



Large Language Models

LLM Tasks

- LLMs can cover all NLP-tasks
- Text Generation
 - Writing assistance, story generation

Translation

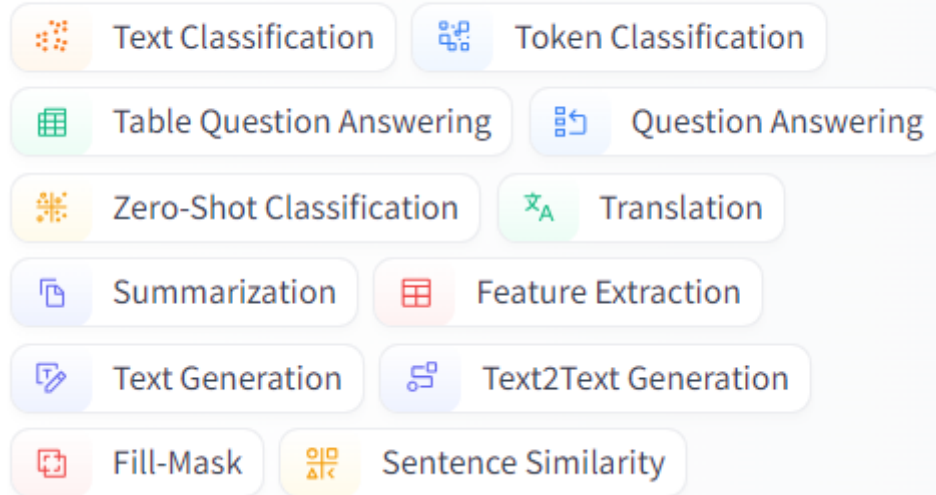
Conversational Agents

Chatbots, virtual assistants

Text summarization

...

Natural Language Processing



Source: <https://huggingface.co/models>

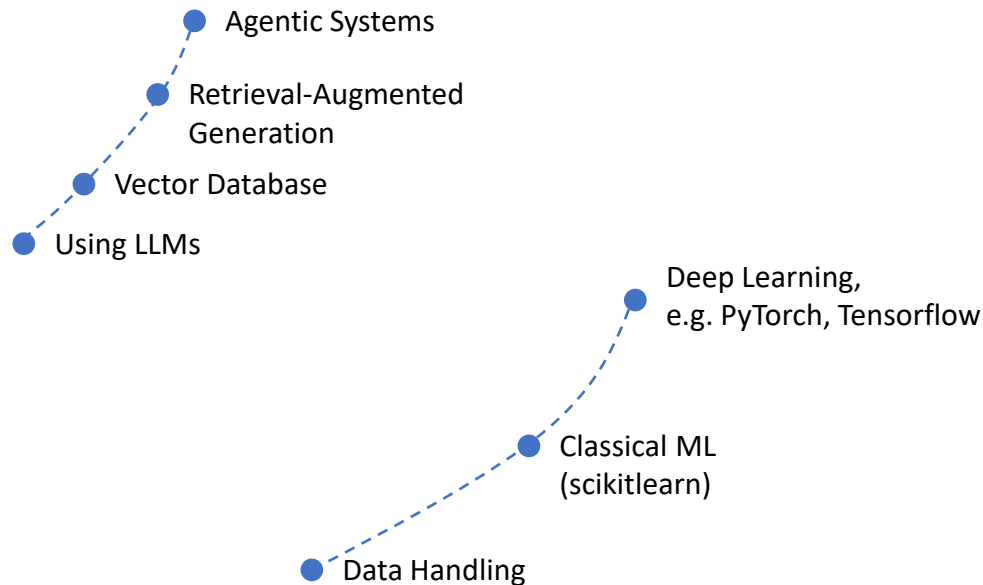


gollnickdata.de

Large Language Models

Model Performance, more Capabilities

Performance /
Capabilities



gollnickdata.de

Large Language Models

Available Providers & Models



- GPT-4o
- GPT-4o mini
- o1-preview / mini
- GPT-4 (Turbo)
- GPT-3.5 Turbo



- Gemini-1.5 Pro
- Gemini-1.5 Flash



- Grok-2

ANTHROPIC

- Claude 3.5 Sonnet

Proprietary /
closed source



- Llama 3.1 family



- Mistral 8x7b

open weight



gollnickdata.de

Large Language Models

LLM Benchmarks

[Arena \(battle\)](#) [Arena \(side-by-side\)](#) [Direct Chat](#) [Leaderboard](#) [About Us](#)

🏆 Chatbot Arena LLM Leaderboard: Community-driven Evaluation for Best LLM and AI chatbots

[Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#) | [Kaggle Competition](#)

Chatbot Arena ([lmarena.ai](#)) is an open-source platform for evaluating AI through human preference, developed by researchers at UC Berkeley [SkyLab](#) and [LMSYS](#). With over 1,000,000 user votes, the platform ranks best LLM and AI chatbots using the Bradley-Terry model to generate live leaderboards. For technical details, check out our [paper](#).

Chatbot Arena thrives on community engagement — cast your vote to help improve AI evaluation!

New Launch! Jailbreak models at [RedTeam Arena](#).

Arena

NEW: Overview

Arena (Vision)

Arena-Hard-Auto

Full Leaderboard

Total #models: 145. Total #votes: 1,898,013. Last updated: 2024-09-17.

Code to recreate leaderboard tables and plots in this [notebook](#). You can contribute your vote at [lmarena.ai](#)

Category

Overall

Apply filter

☐ Style Control

☐ Show Deprecate

Overall Questions

#models: 145 (100%) #votes: 1,898,013 (100%)

Rank* (UB)	Model	Arena Score	95% CI	Votes	Organization	License	Knowledge Cutoff
1	o1-preview	1355	+12/-11	2991	OpenAI	Proprietary	2023/10
2	ChatGPT-4o-latest... (2024-09-03)	1335	+5/-6	10213	OpenAI	Proprietary	2023/10
2	o1-mini	1324	+12/-9	3009	OpenAI	Proprietary	2023/10
4	Gemini-1.5-Pro-Exp-0827	1299	+5/-4	28229	Google	Proprietary	2023/11
4	Grok-2-08-13	1294	+4/-4	23999	xAI	Proprietary	2024/3
6	GPT-4o-2024-05-13	1285	+3/-3	98695	OpenAI	Proprietary	2023/10
7	GPT-4o-mini-2024-07-18	1273	+3/-3	30434	OpenAI	Proprietary	2023/10

Source: <https://lmarena.ai/>



gollnickdata.de

Large Language Models

Practical Coding: First LLM Interaction

1.

API Key Setup



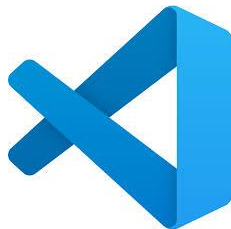
2.

Package Installation



3.

LLM Use
Python Script



gollnickdata.de

Large Language Models

Message Types

System Message

- defines how the model should react
- works like role-play
- Example: „You are a helpful AI assistant designed to provide accurate, concise, and polite responses“

User Message

- user input

AI Message

- corresponds to model response
- different properties,
- mainly „content“ relevant
- more information on input and output tokens available, ...



Large Language Models

Model Selection



Price



On-Prem vs. Cloud



Performance



Open Source vs.
Open Weight



Knowledge-Cutoff



Context-Window



Latency



gollnickdata.de

Large Language Models

Introduction

Artificial Narrow Intelligence (ANI)

- Designed for a specific task
- Limited to scope to well-defined task-specific applications

Artificial General Linguistic Intelligence (AGLI)

- Advanced general capabilities specifically in language understanding and generation
- Examples: GPT-4, Claude, Gemini, Llama, Mistral

Artificial General Intelligence (AGI)

- AI systems with ability to understand, learn, and apply knowledge across broad range of tasks
- Targets all cognitive tasks, generalize knowledge

