

O'REILLY®

# How Do LLMs Work?

Bert Gollnick





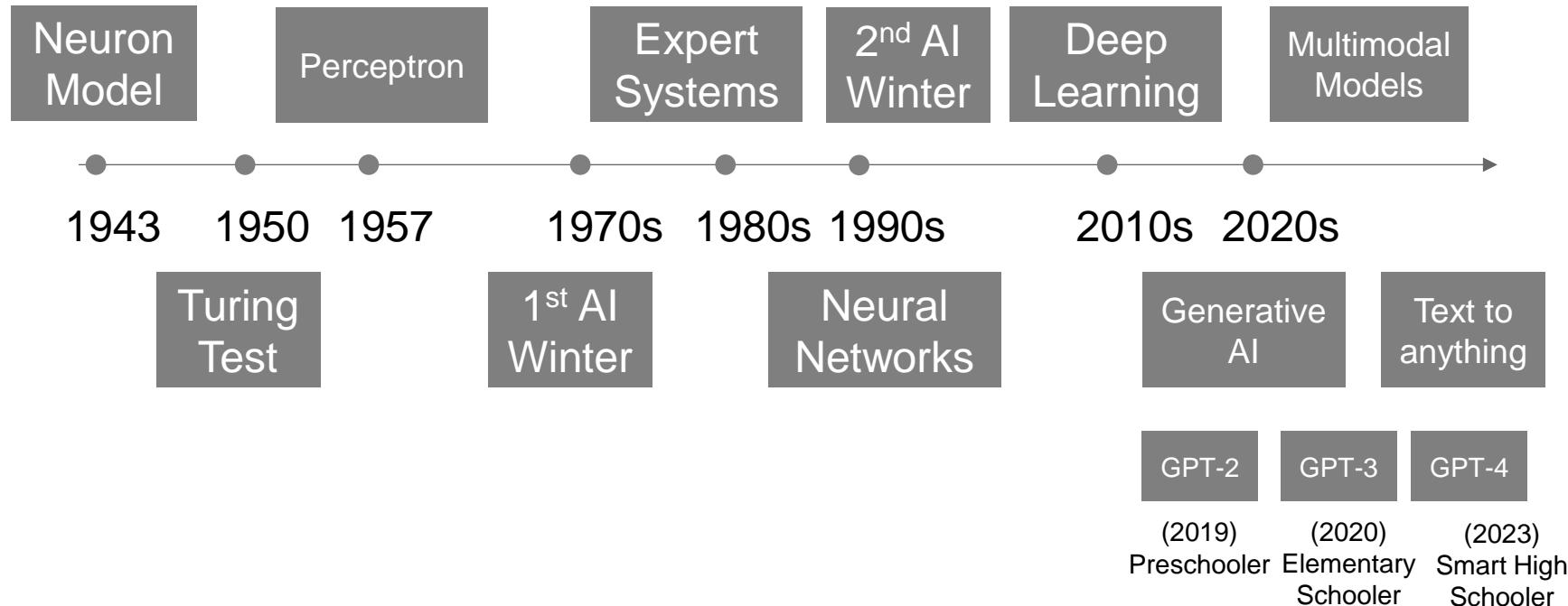
# Learning Objectives

**By the end of this module, you will:**

- Know what LLMs are
- Understand how they work
- Explore some of their applications
- Recall the evolution of AI (until now)

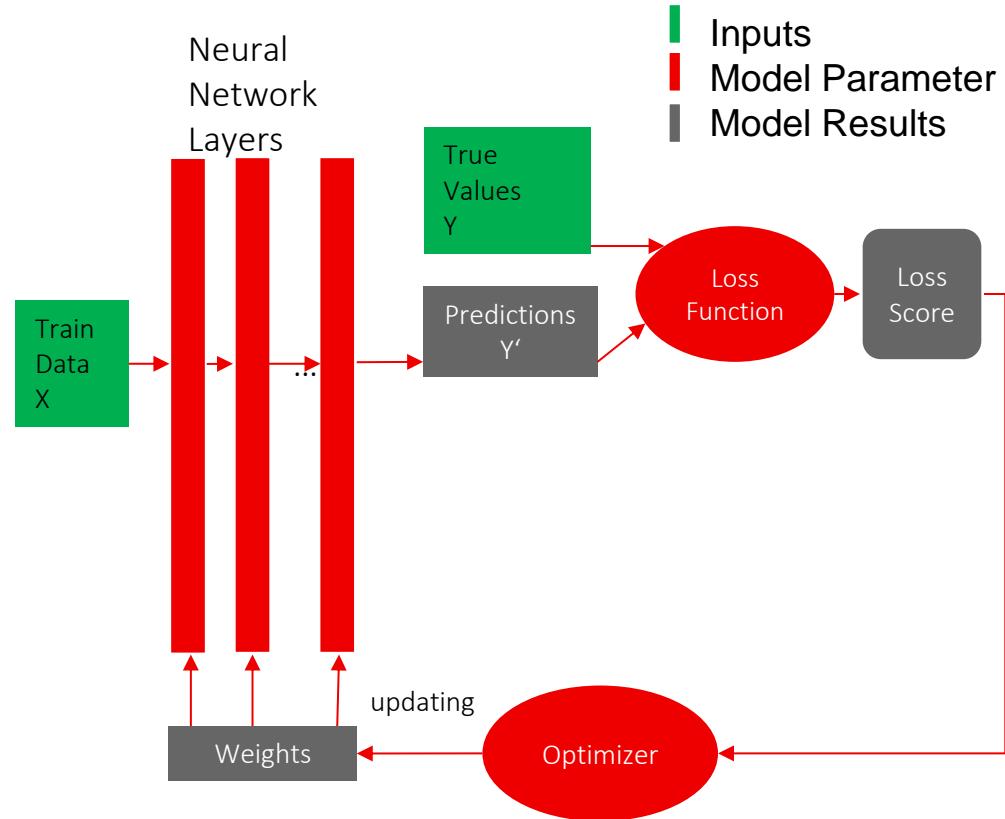


# Evolution



# Deep Learning

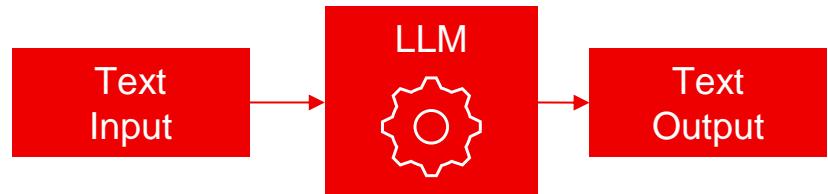
- Subset of machine learning using neural networks with multiple layers
- Capable of learning complex patterns in data
- Effective in complex tasks like Computer Vision and Natural Language Processing





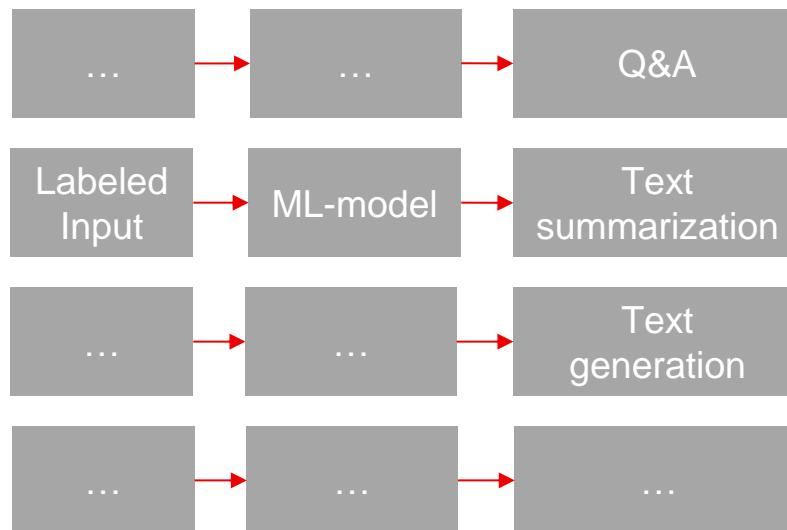
# What are Large-Language Models?

- 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

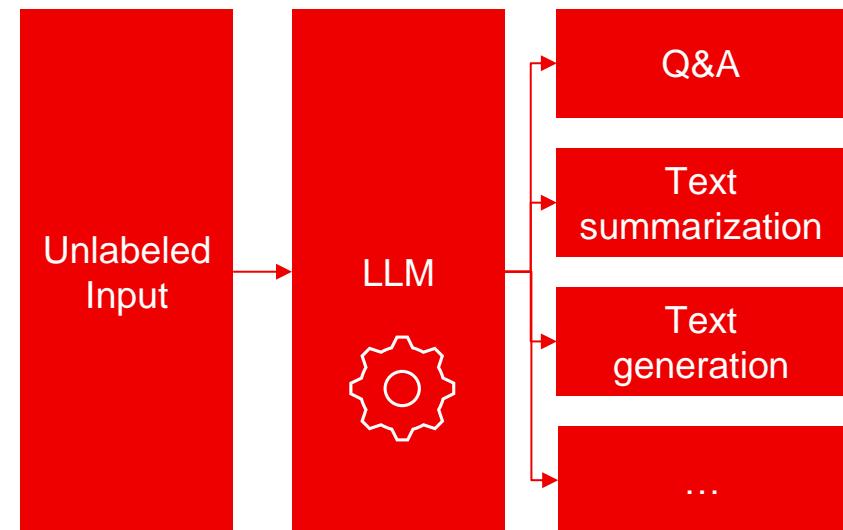




# Difference to Classical Models



Classical ML-models



Classical ML-models



# Transformers

- Type of neural network architecture designed for handling sequential data, primarily used in Natural Language Processing (NLP) tasks
- Huge performance increase over other architectures
- Foundation for all Large-Language Models
- Applied also in Computer Vision, or Audio Processing
- Originated from 2017 paper „Attention is all you need“ (Vaswani et al.)
- Applies Encoder-Decoder structure
- Variants: BERT, GPT



# Applications

- LLMs can cover all NLP-tasks
- Text Generation
  - Writing assistance, story generation
- Translation
- Conversational Agents
  - Chatbots, virtual assistants
- Text summarization
- ...

## Natural Language Processing

 Text Classification

 Token Classification

 Table Question Answering

 Question Answering

 Zero-Shot Classification

 Translation

 Summarization

 Feature Extraction

 Text Generation

 Text2Text Generation

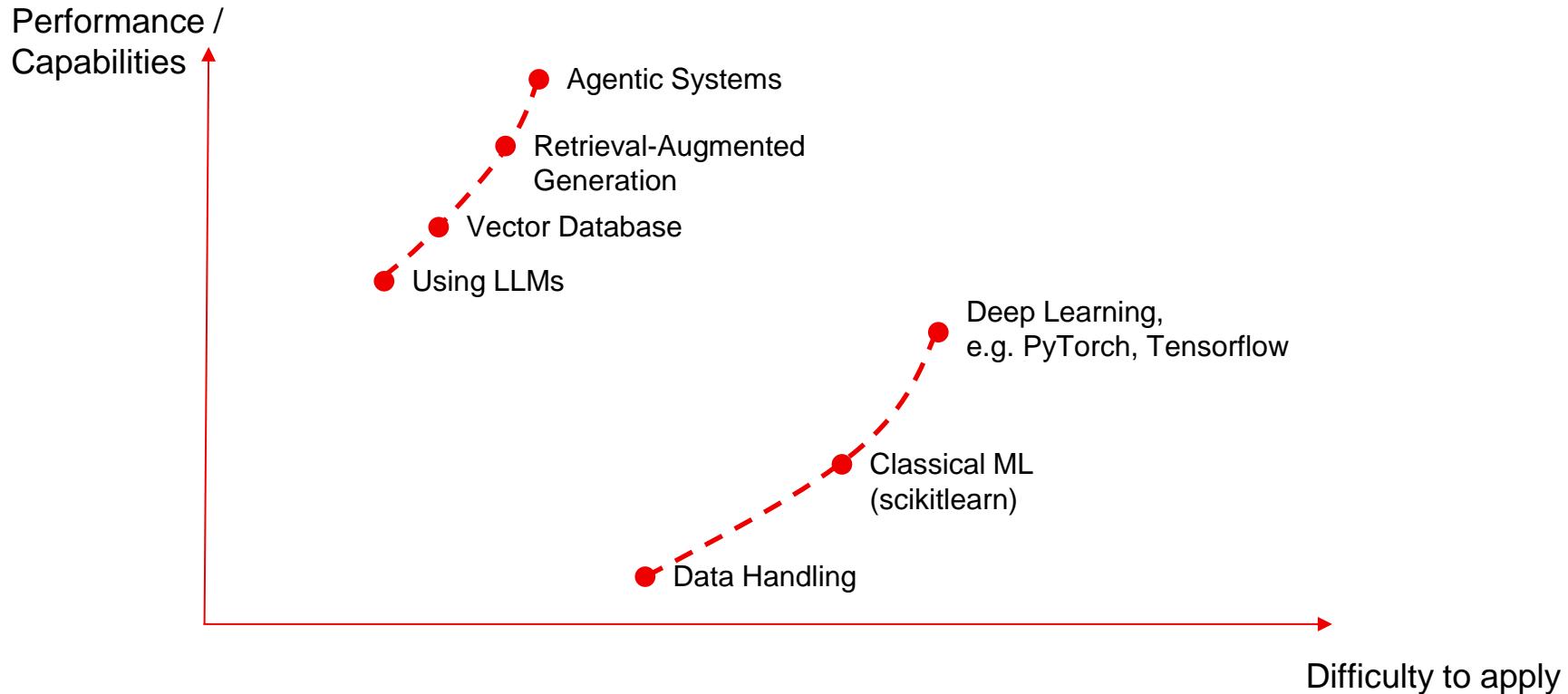
 Fill-Mask

 Sentence Similarity

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



# More performance, easier use





# ANI, AGLI, and AGI

## 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

The background features a vibrant, warm color gradient transitioning from deep red on the left to bright yellow on the right. Overlaid on this gradient are several large, semi-transparent circular shapes in shades of red, orange, and yellow, creating a layered, sunburst-like effect.

O'REILLY®