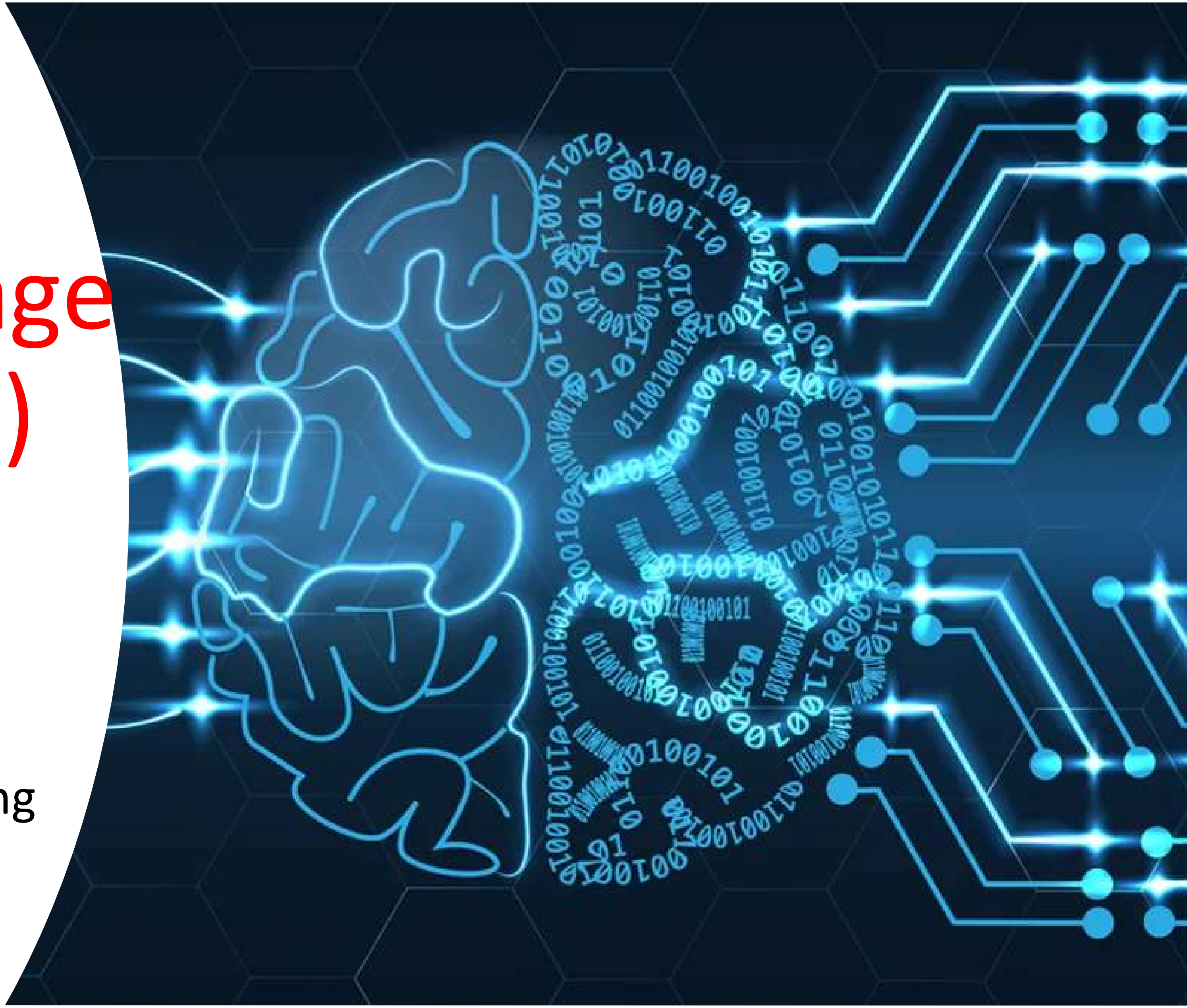
Large Language Models (LLM)

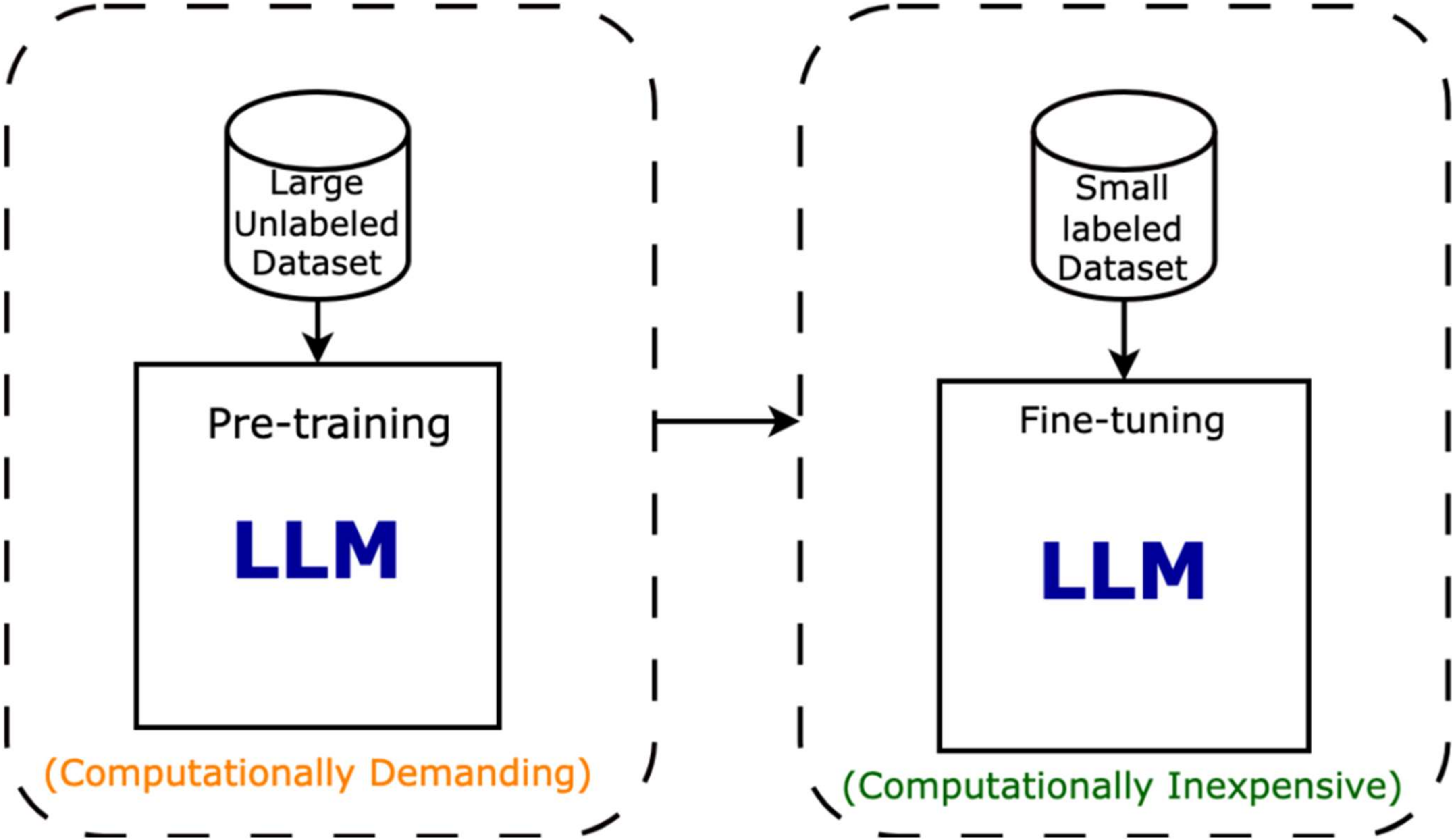
- Module 1
- Lecture 3

Pretraining vs. Fine-Tuning Paradigms



Understanding the Two-Phase Training Process

- Modern LLMs are trained in two stages:
  - Pretraining: Unsupervised learning on massive text corpora.
  - Fine-Tuning: Supervised or instruction-based adjustment on task-specific data.
- This paradigm helps build general-purpose models adaptable to many domains.


Large Unlabeled Dataset
Pre-training
LLM
(Computationally Demanding)
Small labeled Dataset
Fine-tuning
LLM
(Computationally Inexpensive)

The Workflow

Pretraining → Transfer → Fine-tuning

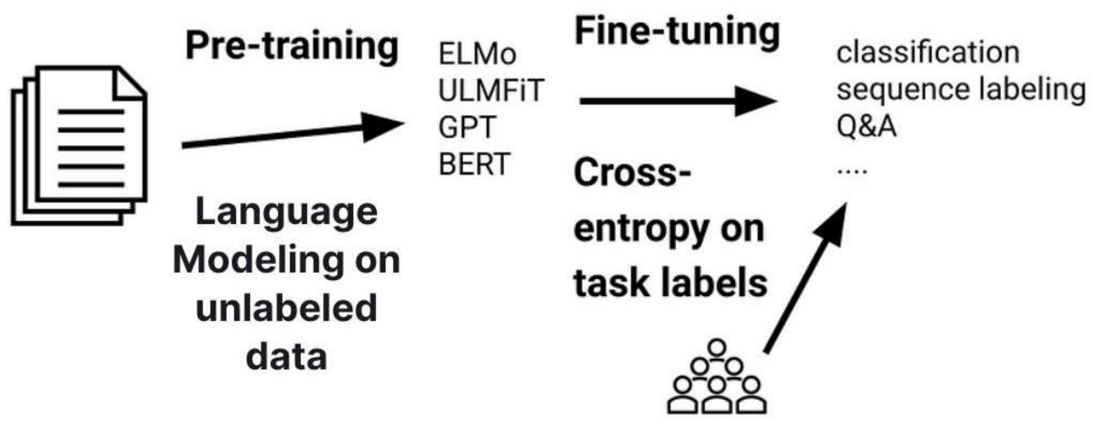
- Step 1: Pre-trained
  - Starting point where model develops strong understanding of language for solving a task.
- Step 2: Transfer
  - Representations learned during pre-trained are transferred and adapted to the new tasks.
- Step 3: Fine tuning
  - Updates the model's parameters to align them with the requirements of the specific tasks.

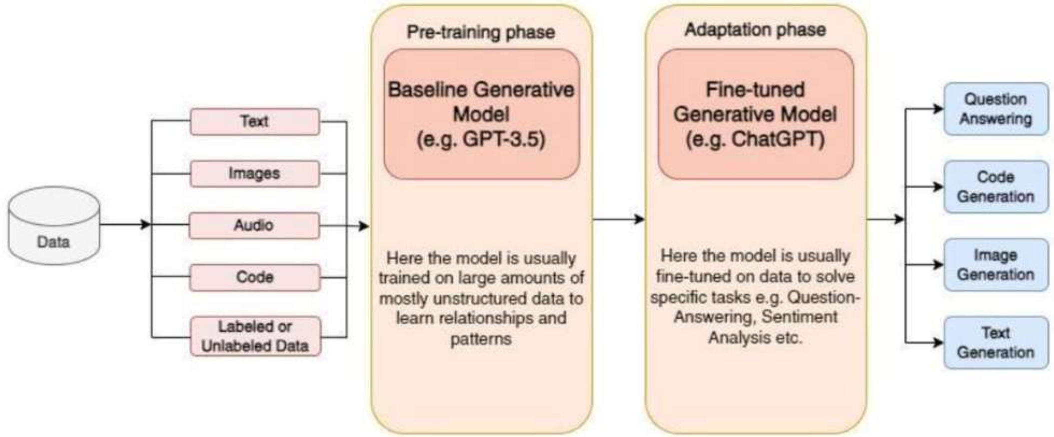
Stage 1: Pretraining Explained

- Pretraining uses self-supervised learning on web-scale corpora (e.g., Wikipedia, books).
- No labels required—models learn by predicting masked/missing tokens.
- Common Objectives:
    - Masked Language Modeling (MLM) → BERT
    - Causal Language Modeling (CLM) → GPT
- Builds foundational understanding of grammar, semantics, and factual knowledge.

Stage 2: Fine-Tuning Explained

- Fine-tuning adjusts the pretrained model for a specific downstream task.
- Requires labeled data (QA pairs, classification labels, etc.).
- Examples:
  - Sentiment Analysis: Classify tweets as positive/negative.
  - Question Answering: Match answers to SQuAD-style inputs.
  - Instruction-Tuning: Use prompts and human feedback (e.g., InstructGPT).


Pre-training
Language Modeling on unlabeled data
ELMo
ULMFiT
GPT
BERT
Fine-tuning
Cross-entropy on task labels
classification
sequence labeling
Q&A
....


Data
Text
Images
Audio
Code
Labeled or Unlabeled Data
Pre-training phase
Baseline Generative Model (e.g. GPT-3.5)
Here the model is usually trained on large amounts of mostly unstructured data to learn relationships and patterns
Adaptation phase
Fine-tuned Generative Model (e.g. ChatGPT)
Here the model is usually fine-tuned on data to solve specific tasks e.g. Question-Answering, Sentiment Analysis etc.
Question Answering
Code Generation
Image Generation
Text Generation

Comparison: Pretraining vs. Fine-Tuning

- Pretraining:
  - Task: General language modeling
  - Data: Unlabeled large corpus
  - Goal: Learn syntax, semantics

- Fine-Tuning:
  - Task: Task-specific learning
  - Data: Labeled or prompted
  - Goal: Optimize for specific tasks

Example: BERT

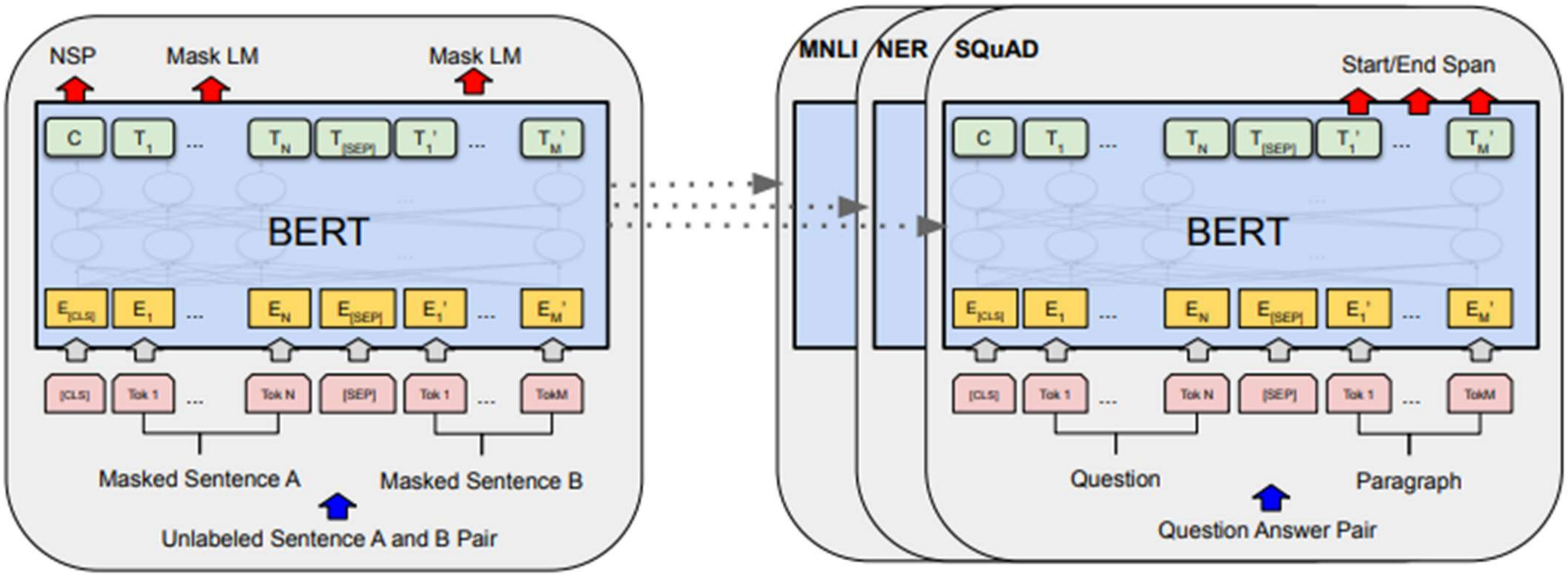
- Pretraining:

- Masked Language Modeling: Predict [MASK] tokens.
- Next Sentence Prediction: Understand sentence relationships.

- Fine-Tuning Tasks:

- Named Entity Recognition (NER)
- Question Answering (SQuAD)
- Sentiment Classification

- Adds task-specific layers to the pretrained transformer encoder.



Pre-training

Fine-Tuning

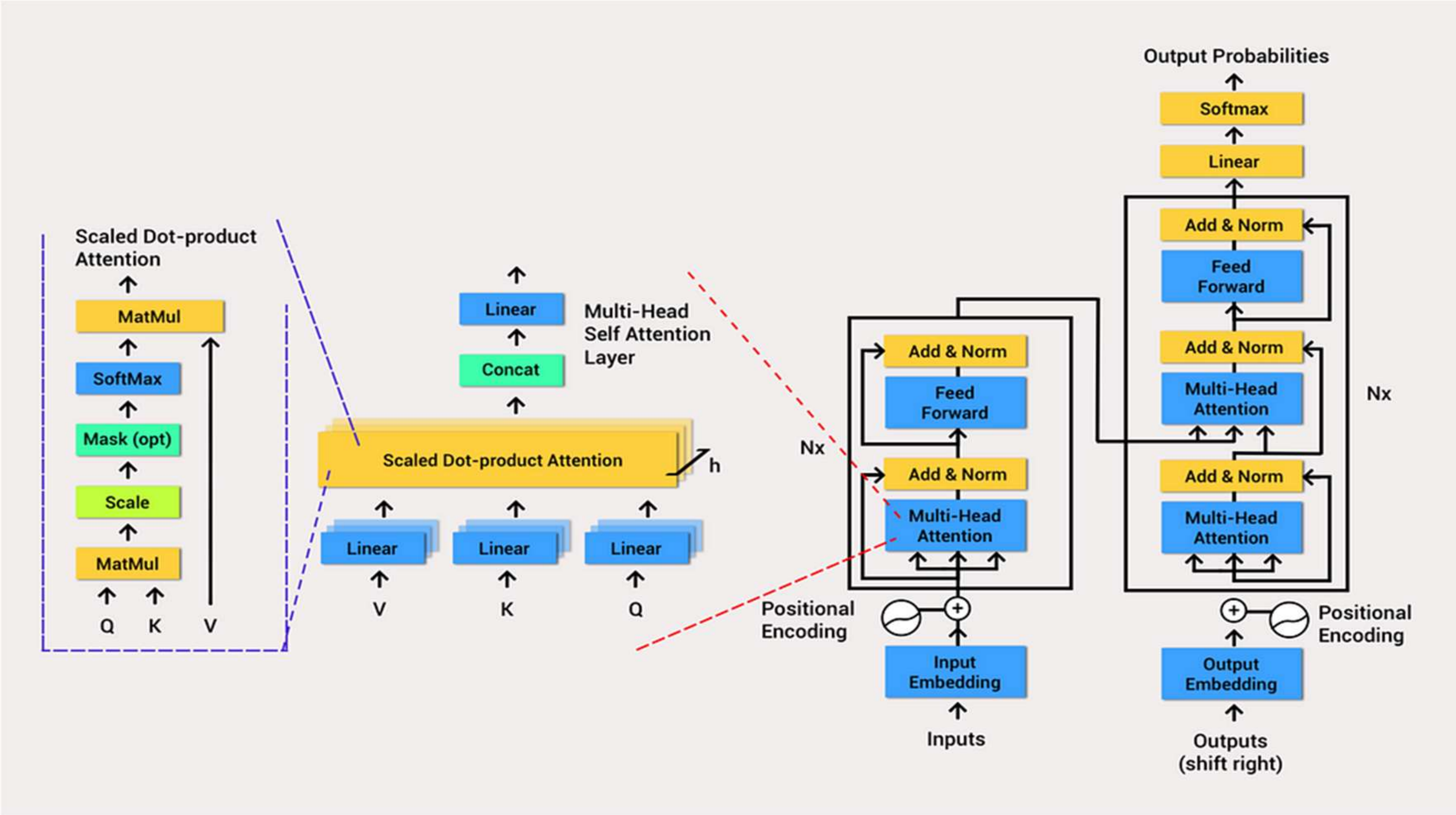
Example: GPT

- Pretraining: Causal language modeling.

- Fine-Tuning examples:

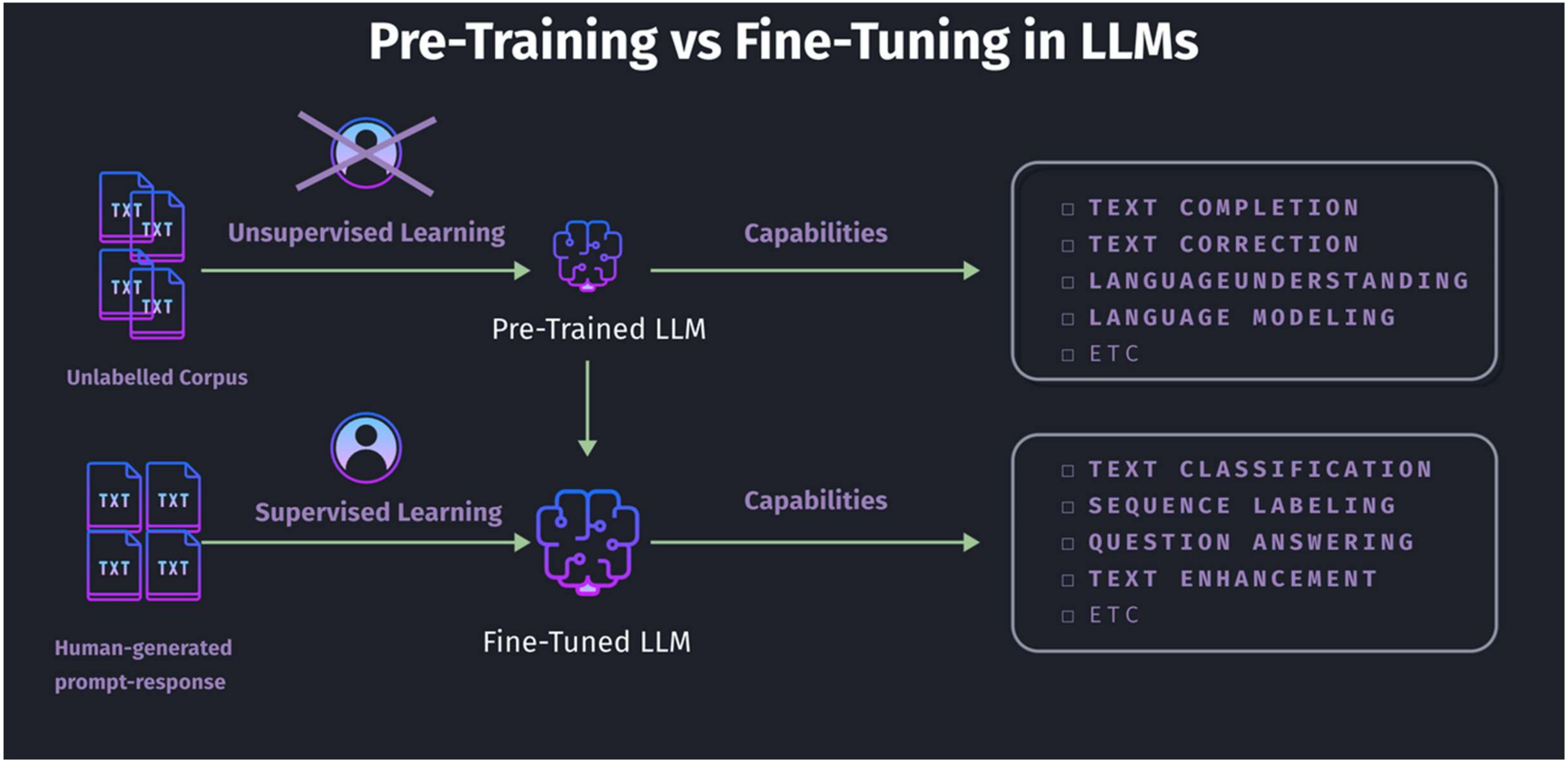
- GPT-1/2/3 Pretraining: Autoregressive (Causal) Language Modeling.
- GPT-3: Few-shot and zero-shot via prompting (no weight update).
- InstructGPT: Fine-tuned with RLHF (human feedback + reward model).
- ChatGPT/GPT-4: Continuation of this with instruction-following and safety alignment.

Architecture of GPT



Pretraining vs. Fine-Tuning: Key Differences

Aspect	Pretraining	Fine-Tuning
Data	Unlabeled web-scale corpora	Labeled task-specific sets
Goal	General language learning	Specialize for specific tasks
Example Models	GPT, BERT	InstructGPT, T5 fine-tuned
Cost	High compute/time	Lower, but requires labels
Flexibility	Very broad/general	Narrow but targeted


Pre-Training vs Fine-Tuning in LLMs
TXT
TXT
TXT
TXT
Unlabelled Corpus
Unsupervised Learning
Pre-Trained LLM
Capabilities
TEXT COMPLETION
TEXT CORRECTION
LANGUAGEUNDERSTANDING
LANGUAGE MODELING
ETC
TXT
TXT
TXT
TXT
Human-generated
prompt-response
Supervised Learning
Fine-Tuned LLM
Capabilities
TEXT CLASSIFICATION
SEQUENCE LABELING
QUESTION ANSWERING
TEXT ENHANCEMENT
ETC

Pre-Training Features	Fine-Tuning Features
Broad language understanding	Task-specific adaptation
Large, diverse dataset training	Smaller, targeted dataset training
General knowledge base development	Rapid specialization
Facilitates transfer learning	Quick learning from few examples
High initial computational cost	Lower computational cost
Scalable with continual learning	Customizable to current data
Sets performance benchmarks	Enhances specific task performance
Flexible across various applications	Efficient for niche applications

References

- https://ankushmulkar.medium.com/what-is-chat-gpt-how-could-its-creators-benefit-from-chat-gpt-c4f4bec80575
- https://medium.com/mantisnlp/supervised-fine-tuning-customizing-llms-a2c1edbf22c3
- https://medium.com/@eordaxd/fine-tuning-vs-pre-training-651d05186faf
- https://arxiv.org/pdf/2402.06853
- https://link.springer.com/content/pdf/10.1007/s10462-024-10888-y.pdf