Lecture-6 : Stages of building LLMS
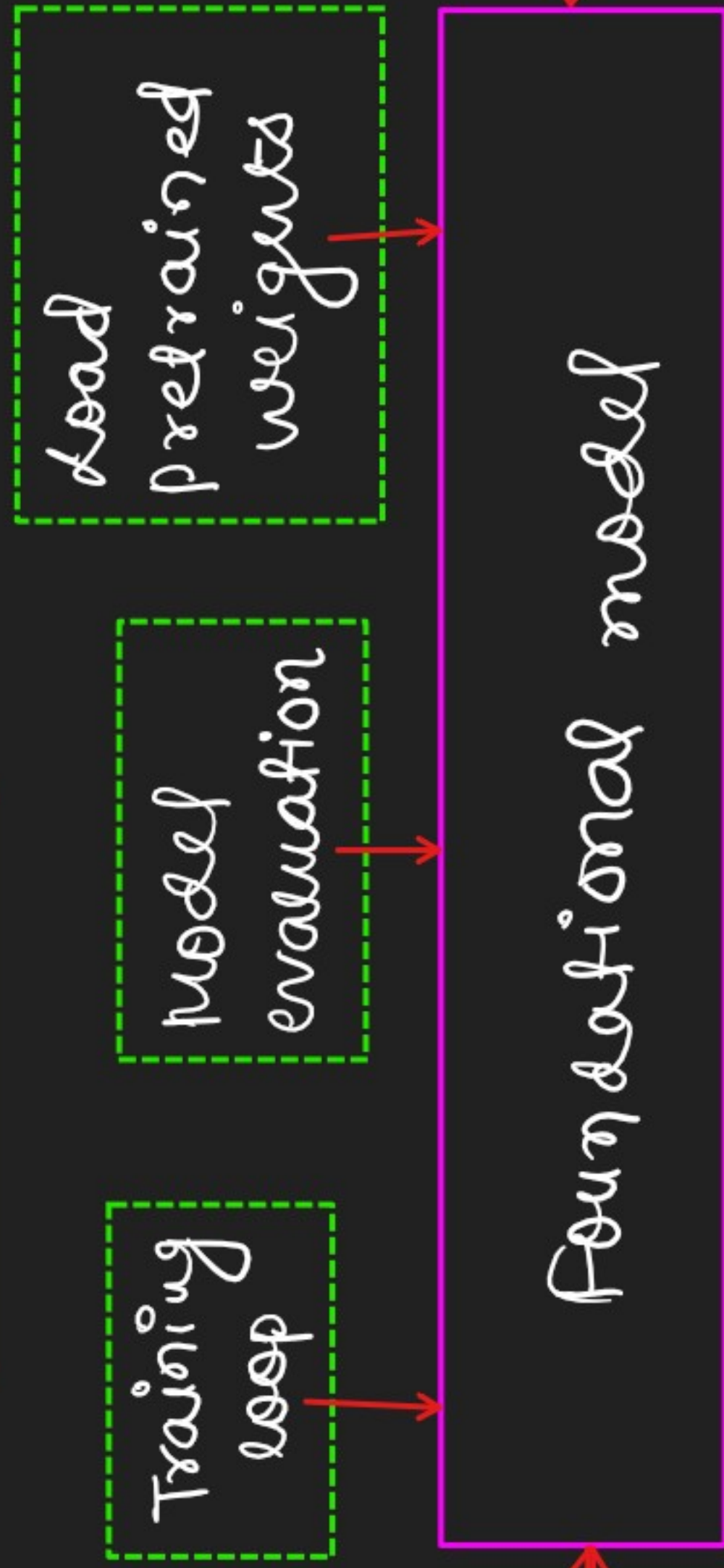
STAGE - 1

Building an LLM

- Data preparation and sampling
- Attention mechanism
- LLM architecture
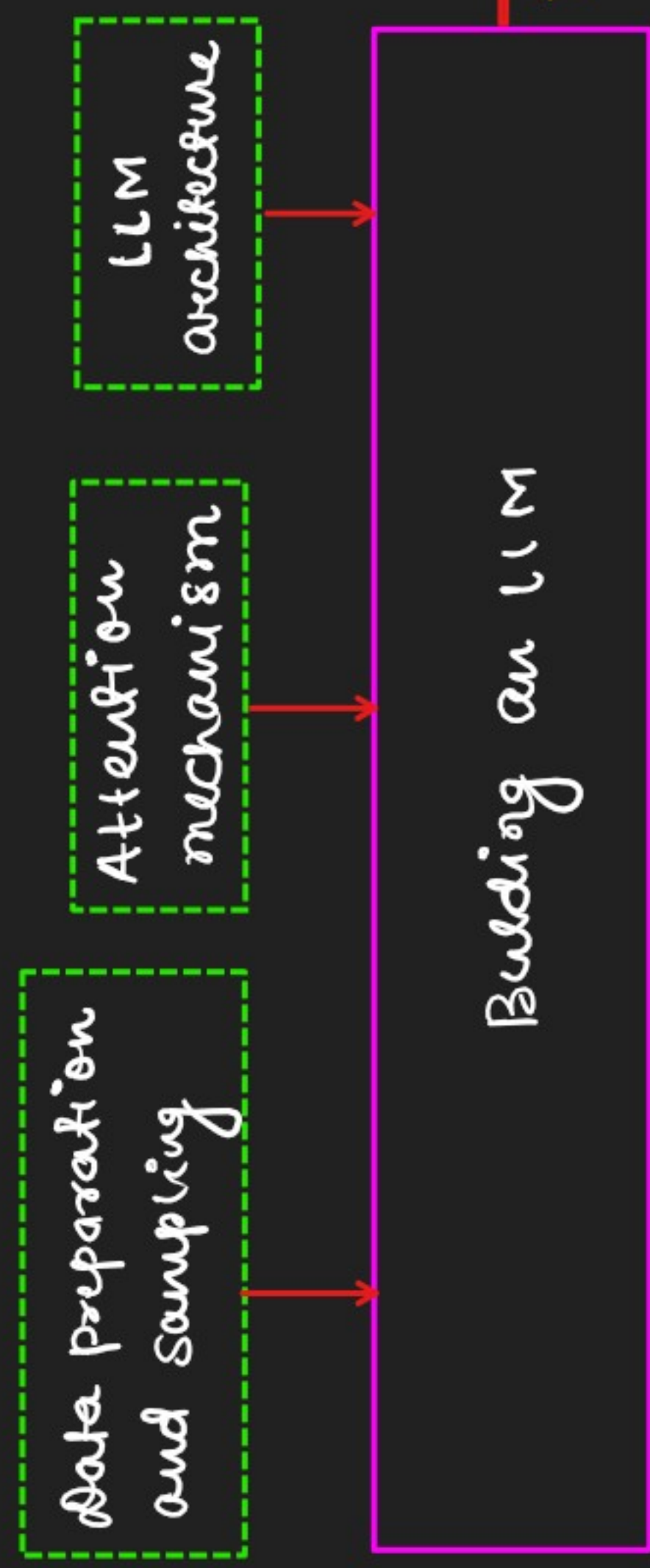
Implement data sampling: understand basic mechanism

→ **Pre-training**

STAGE - 2

Foundational model

- Training loop
- Model evaluation
- Load pretrained weights

Pretrain an LLM on unlabeled data.

STAGE - 3

Fine tuning → Classifier

Fine tuning → Personal Assistant
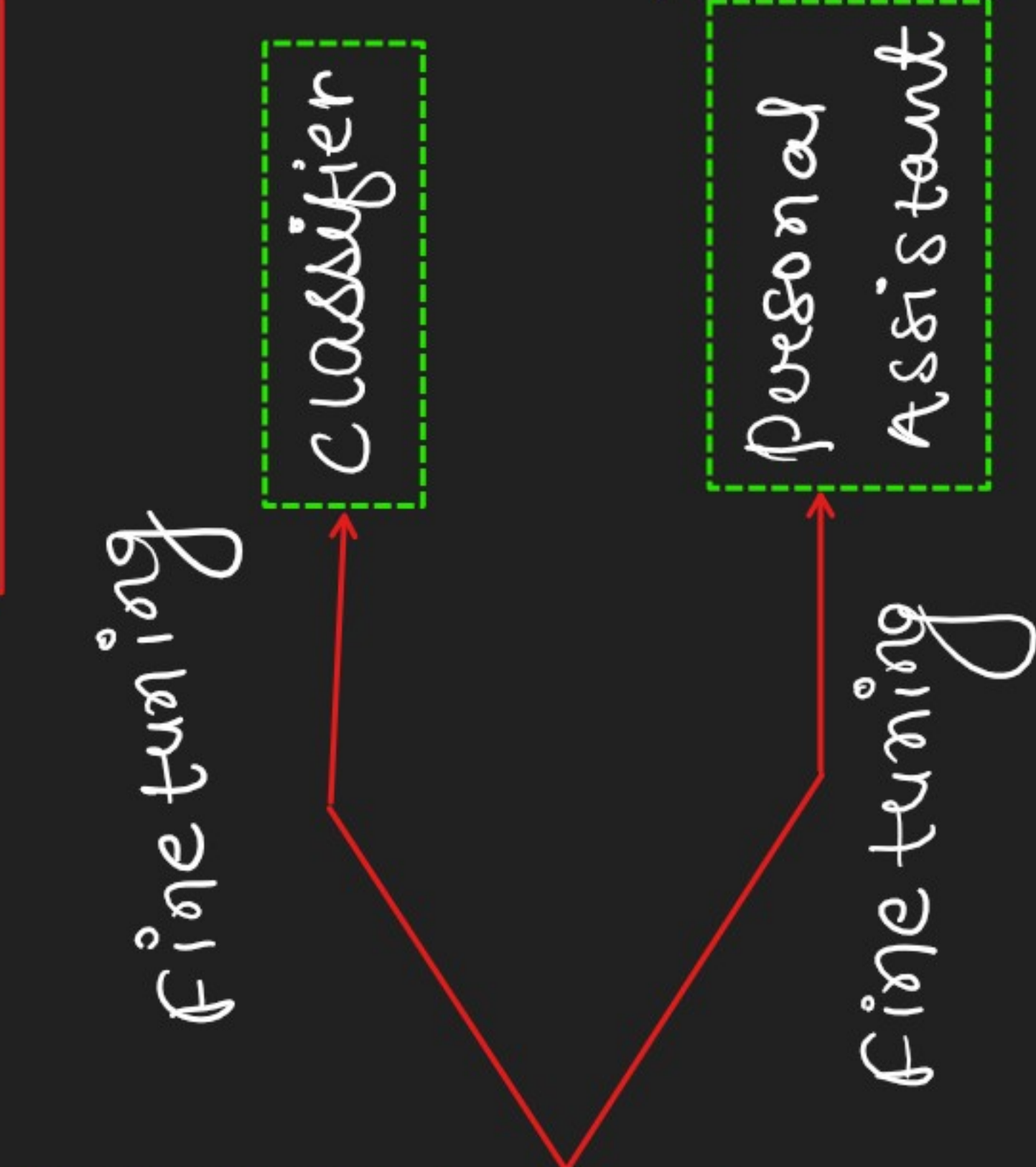
Finetune the pretrained LLM

Recap of what all we have learnt so far:~

1. LLMs have transferred the field of NLP. They have led to advancement in generating, understading and translating human language.

2. Modern LLMs are trained in 2 main steps:-

Pretrained on unlabelled data.
(foundational data)

Finetuning on a smaller, labelled dataset eg. spam, not spam.

very large datasets needed (billions of words)

finetune LLMS can outperform just pretrained LLMs on specific tasks.

It cost a lots!

3. LLMs are based on transformer architecture.

↓

Key idea - Attention mechanism gives LLM selective access to whole input sequence when generating output one word at a time.

4. Original transformer = encoder + Decoder

5. GPT: only has Decoder, no encoder

6. while LLMs are only trained for predict next word, they show emergent property.

↳ ability to classify, translate, and summarize texts.