⭐ Verifier Engineering ⭐

Search → **Verify** → **Feedback** | **Examples**

Search

Search
● :token/sentence

Q → ● → ●
Linear

Q → ● ● ●
Tree

Candidate Response

Verifier

Search Engine Human Code interpreter

Data Base Rules Reward Model

Verifier Combinations

(nearly) Golden Verification

Feedback

I. Training-based

SFT | DPO | PPO

II. Inference-based

Q → 1 → 2 → 3 **Verifier Aware**

Q → 1 → 2 → 3 **Verifier Guided**

Examples

RLHF

Q → ● → ●
Linear Reward Model **PPO**

OmegaPRM

Q → ● ● ●
Tree Reward Model **Verifier Guided**

Experiential Co-Learning

Q → ● → ●
Linear LLMs **Verifier Aware**