

Techniques for Developing and Refining Datasets

1. Data Collection and Curation
2. Cleaning and Preprocessing
3. Annotation and Labeling
4. Validation and Testing

Comparison of Language Model Fine-Tuning Approaches

1. Full Fine-Tuning

Fine-tune all layers of the pre-trained model on domain-specific data.

2. Parameter-Efficient Fine-Tuning (e.g., LoRA, Prefix Tuning)

Modify or add a small number of parameters (adapters) to the model, leaving the core weights unchanged.

3. Prompt Tuning

Optimize a set of task-specific prompts that guide the model's response without altering its weights.

4. Zero-shot and Few-shot Learning

Use the model's pre-trained capabilities with minimal or no task-specific training.