Fine-tuning Large Language Models with Limited Data: A Survey and Practical Guide

Marton Szep[1,2] **Daniel Rueckert**[1,3,4] **Rüdiger von Eisenhart-Rothe**[2]
Florian Hinterwimmer[1,2]

[1]AI in Healthcare and Medicine, Technical University of Munich (TUM) and
TUM University Hospital, Munich, Germany
[2]Department of Orthopaedics and Sports Orthopaedics, TUM University Hospital, Germany
[3]Munich Center for Machine Learning (MCML), Munich, Germany
[4]Department of Computing, Imperial College London, UK
marton.szep@tum.de

Abstract

Fine-tuning large language models (LLMs) with limited data poses a practical challenge in low-resource languages, specialized domains, and constrained deployment settings. While pre-trained LLMs provide strong foundations, effective adaptation under data scarcity requires focused and efficient fine-tuning techniques. This paper presents a structured and practical survey of recent methods for fine-tuning LLMs in data-scarce scenarios. We systematically review parameter-efficient fine-tuning techniques that lower training and deployment costs, domain and cross-lingual adaptation methods for both encoder and decoder models, and model specialization strategies. We further examine preference alignment approaches that guide model behavior using limited human or synthetic feedback, emphasizing sample and compute efficiency. Throughout, we highlight empirical trade-offs, selection criteria, and best practices for choosing suitable techniques based on task constraints, including model scaling, data scaling, and the mitigation of catastrophic forgetting. The aim is to equip researchers and practitioners with actionable insights for effectively fine-tuning LLMs when data and resources are limited.

1 Introduction

Pre-trained language models (PLMs) have driven unprecedented advances in NLP, showing strong capabilities across a wide range of tasks (Han et al., 2021). However, their development and adaptation typically require large-scale data and computational resources, often unavailable in real-world scenarios, particularly in low-resource languages or specialized domains such as medicine (Crema et al., 2023), law (Noguti et al., 2023), chemistry (Jablonka et al., 2024), or finance (Zhao et al., 2021). To address these constraints, transfer learning has become the standard paradigm: Models are first pre-trained on general corpora and then adapted through fine-tuning (FT) and preference alignment (PA). Yet, these adaptation stages can still demand significant annotated data, and naive FT under low-data conditions risks overfitting, catastrophic forgetting (Ramasesh et al., 2021), and poor generalization.

Given the rapidly evolving field and broad array of methods, a structured and pragmatic review is necessary to navigate current strategies and make informed choices under real-world constraints. This paper offers a practical guide to fine-tuning large language models under data scarcity. We examine recent developments across four key areas. First, we review **(1) Parameter-efficient Fine-tuning** (PEFT, § 3), which enables modular and cost-effective adaptation. Second, we explore **(2) Domain and Cross-lingual FT** (§ 4), targeting robustness to distributional shifts. Third, **(3) FT for Specialization** (§ 5) covers techniques for specialized tasks, domains, and low-resource languages, including scaling, data efficiency, and design recommendations. Fourth, we analyze **(4) Preference Alignment** (PA § 6), including data requirements, scaling behavior, and downstream effects. We conclude with a focused discussion on practical challenges and emerging strategies, such as the trade-offs between encoder and decoder architectures, cross-lingual transfer in low-resource contexts, and the potential of model merging. Our emphasis throughout is on actionable insights and contextualized, concrete

Transactions of the Association for Computational Linguistics, vol. 14, pp. 341–377, 2026. https://doi.org/10.1162/TACL.a.627
Action Editor: Sebastian Ruder. Submission batch: 11/2024; Revision batch: 6/2025; Published 4/2026.
© 2026 Association for Computational Linguistics. Distributed under a CC-BY 4.0 license.

guidance to support effective LLM fine-tuning when data is limited.

2 Related Surveys

Several prior works address related aspects or have partial overlap with our work. Some surveys target data augmentation as a solution to data scarcity (Feng et al., 2021; Chen et al., 2023b), including domain-specific generation techniques for low-resource settings (Stylianou et al., 2023). Model-centric strategies are covered in surveys on PEFT methods (Han et al., 2024; Lialin et al., 2024), while Treviso et al. (2023), Wan et al. (2024), and Xu et al. (2025) offer broader coverage of efficient training techniques across different stages of model development. Other works focus on low-resource supervised NLP (Hedderich et al., 2021), instruction tuning (Zhang et al., 2024), and PA (Gao et al., 2024; Jiang et al., 2024). In contrast, our work offers a pragmatic overview of the full post-pretraining pipeline for adapting PLMs with limited data. We cover PEFT, domain and cross-lingual generalization, specialization strategies, and PA in an integrated survey, emphasizing practical insights, comparative trade-offs, and methodological guidance to support informed decisions in data-scarce scenarios.

3 Parameter-efficient Fine-tuning

Fine-tuning (FT) all parameters of large PLMs is computationally expensive, sample-inefficient, and often unstable in low-resource regimes (Dodge et al., 2020). Parameter-efficient fine-tuning (PEFT) mitigates these issues by updating only a small fraction of weights, enabling strong adaptation with lower cost and reduced risk of catastrophic forgetting in data-scarce scenarios. We organize PEFT methods by how they *compose* with the base model parameters (following Pfeiffer et al., 2023), and highlight key performance and efficiency considerations for selecting the appropriate method for a given task and data setup. For comprehensive surveys, see Han et al. (2024) and Pfeiffer et al. (2023).

3.1 Parameter Composition

These methods modify existing weights through sparse or low-rank updates.

Selective methods train only a subset of the model weights (specific layers, parameter types, etc.) while keeping the rest fixed. They build on the inductive bias of sparsity: In an over-parameterized model, a small fraction of parameters plays a disproportionately important role for a particular task. Variants include tuning specific (usually last) layers or model components like the bias, or LayerNorm parameters (Devlin et al., 2019; Ben Zaken et al., 2022; Liu et al., 2022). Another line of work optimizes subnetworks based on gradient information either by selecting a specific subnetwork (Xu et al., 2021b; Ansell et al., 2022; Sung et al., 2021), or adaptively choosing it through a bi-level optimization strategy (Zhang et al., 2022a; Yu et al., 2023). These methods are extremely lightweight and overhead-free at inference time, but usually underperform compared to approaches that introduce new trainable parameters.

Reparametrization methods replace full weight updates with low-dimensional structures, motivated by the low-rank nature of PLM weights (Aghajanyan et al., 2021). The most widely used approach, LoRA (Hu et al., 2021), decomposes weight updates into the product of two low-rank matrices and can be merged into the backbone at inference, avoiding extra latency. Its effectiveness led to numerous extensions improving efficiency, robustness to hyperparameters, and calibration (Liu et al., 2024b; Dettmers et al., 2023; Hayou et al., 2024; Ding et al., 2023a; Zhang et al., 2022b; Edalati et al., 2022). Reparametrization methods work best on larger models and with moderate to high data sizes, and often match full FT performance with careful tuning (Van Veen et al., 2023).

3.2 Input Composition

These methods extend the input with learned embeddings that act as task-specific context.

Soft prompt prepend learned embedding vectors to the input or key-value pairs in attention layers (Li and Liang, 2021; Lester et al., 2021). They are a popular choice to adapt models across domains and languages due to their flexibility and efficiency (Tu et al., 2024; Goswami et al., 2023; Zhao et al., 2022), but often converge slowly and perform reliably only at very large model scales (Mao et al., 2022). Variants such as $\texttt{IA}^3$ (Liu et al., 2022) improve efficiency by learning vectors to

scale hidden activations, achieving adapter-level performance with fewer parameters. Soft prompts typically incur additional inference overhead, making them more suitable for scenarios prioritizing training efficiency over runtime cost.

3.3 Function Composition

These methods augment the frozen backbone with lightweight trainable modules.

Adapters are small feedforward modules inserted between Transformer layers (Houlsby et al., 2019; Pfeiffer et al., 2020). They reduce training memory and enable flexible modular composition across tasks (Pfeiffer et al., 2021; Wang et al., 2022; Chronopoulou et al., 2023), but sequential insertion slows down inference by up to 40% (Rücklé et al., 2021). Parallel and Ladder Side-tuning variants mitigate this issue (He et al., 2021; Sung et al., 2022). Adapters can act as bottlenecks, requiring wider layers and more paramaters than other PEFT methods to maintain performance (Hu et al., 2023b). Their efficiency can be improved through low-rank reparametrization, linking them conceptually with methods in § 3.1.

3.4 Hybrid Methods

Hybrid methods combine different forms of composition to leverage complementary strengths. Examples include Compacter (Karimi Mahabadi et al., 2021), which integrates adapters with low-rank decomposition, and UniPELT, which dynamically activates adapters, prefixes, and LoRA (Mao et al., 2022). Such methods can be powerful in very low-resource settings, but are generally more sensitive to hyperparameter tuning (Chen et al., 2022a).

3.5 Selection Criteria

Choosing the right PEFT method depends on task type, model size, and data scale. Full FT becomes favorable only with million-scale datasets (Zhang et al., 2023a), whereas PEFT often matches or outperforms it under 100k samples (Lialin et al., 2024; Ding et al., 2023b; Jukić and Snajder, 2023). In practice, selective and reparametrization methods update less than 1% of parameters, while adapters and soft prompts can require up to several percent (Lialin et al., 2024). Selective methods are simple, efficient, but they usually underperform (Liu et al., 2022), with LayerNorm tuning sometimes found effective (Lialin et al., 2024). Soft prompts can outperform LoRA on the scale of a few thousand samples (Liu et al., 2022; Zhang et al., 2023a; Ding et al., 2023b), but converge slowly and are fragile at smaller model scales. Adapters and LoRA stand out as the most robust choices, reliably matching full FT with minimal hyperparameter tuning in moderate- to large-scale data settings (over 10k samples; Zhang et al., 2023a). Hybrid methods can yield gains in extremely low-resource scenarios but at the cost of added complexity (Chen et al., 2022a; Mao et al., 2022; He et al., 2021).

Inference overhead varies across methods, with adapters imposing the largest slowdown due to added depth, followed by soft prompts (30–40%), reparametrization methods (although they can usually be merged into the model), and selective methods, which incur minimal to no runtime cost (Hu et al., 2021; Lialin et al., 2024). Table 4 and Appendix A offer more practical guidance on selecting PEFT methods across discriminative tasks for encoder models.

3.6 Efficient Design Choices

Efficient design choices in PEFT methods are vital for maximizing performance gains while minimizing resource requirements. Studies show that PEFT methods exhibit better performance and convergence rate as the model size increases, also progressively narrowing the gap between full FT and PEFT (Zhang et al., 2023a; Ding et al., 2023b; Lialin et al., 2024). Intuitively, with increasing data size expressivity should be expanded through more PEFT parameters. However, studies (Zhang et al., 2023a; Ding et al., 2023b) show that the gains of scaling the tunable parameters (e.g., LoRA rank $4 \rightarrow 128$ or prompt-tuning prompt length $50 \rightarrow 600$) are marginal when scaling the data size up to 100k samples, especially compared to scaling the model. With a fixed compute budget, it is preferable to opt for larger models with reduced precision, e.g., 8- or 4-bit quantization (Dettmers et al., 2023).

4 Domain and Cross-lingual Fine-tuning

With the rapid growth of language models in recent years and the vast knowledge accumulated during large-scale pre-training (PT), research has increasingly focused on how to best shape and adapt this knowledge through FT for improved

	Method	Requirements	Advantages	Limitations	Use Cases
Domain & Cross-lingual FT §4	Continued Pre-training (§4.1)	Monolingual, in-domain corpora	Can use any kind of relevant unlabeled corpora	Computationally intensive; Can lead to catastrophic forgetting	Cross-lingual or domain-specific adaptation
	Pattern-exploiting training (§4.2)	Few labeled examples	Better alignment of FT and PT task	Handcrafting; Slow autoregressive decoding	Few-shot NLU tasks with encoders
	Meta-Learning (§4.2)	Few labeled examples in related tasks	Adapts quickly to new tasks	Computationally intensive, potentially complex to implement	Few-shot classification tasks with encoders
	Intermediate fine-tuning (§4.2)	Data similar to final task	Improves convergence and final task performance	Sensitive to the similarity with the final task	Bridge PT and low-resource task (encoder)
	Multi-task learning (§4.2)	Large, diverse set of related task data	Strong generalization to unseen tasks	High computational and data requirements	General-purpose generative models for zero-shot tasks
FT for Specialization §5	Embedding learning (§5.1)	Parallel or non-parallel data	Better transferability and semantic relationships	Static embeddings lack context	Discrepant pre-trained and target domain or language
	Contrastive learning (§5.2)	Paired data	Useful representations, data-efficiency	Sensitive to data pair quality and quantity	Cross-lingual alignment; Classification tasks
	Adversarial learning (§5.2)	Unpaired data	Learns generalized features	Adversarial training is prone to be unstable	Cross-domain and cross-lingual transfer
	Semi-supervised learning (§5.3)	Labeled & unlabeled data	Utilizes unlabeled data, cost-effective	Sensitive to pseudo-label quality or noisy data	Classification and seq2seq low-resource tasks
	Unsupervised learning (§5.3)	Unpaired data	No need for labeled data, scalable	Often worse performance than supervised	Scenarios with no labeled data
	Active learning (§5.3)	Labeled & unlabeled data	Maximizes utility of labeled data	Requires iterative training and data selection	Classification and seq2seq low-resource tasks

Table 1: Overview of the domain, cross-lingual, and specialization fine-tuning approaches discussed in this paper.

generalization across diverse tasks both for encoder (Schick and Schütze, 2021a,b; Vu et al., 2020) and decoder models (Ouyang et al., 2022; Brown et al., 2020). For a concise summary of the methods, see Table 1.

4.1 Continued Pre-training

Continued pre-training (CPT) bridges the gap between general PT and specialized domains or languages using unlabeled data. It reuses the original self-supervised PT objective (or combines different ones for added capabilities like fill-in-the-middle; Bavarian et al., 2022) allowing adaptation without supervised labels (Gururangan et al., 2020), facilitating vocabulary extension and sample-efficient reuse of labeled FT data (Gnehm et al., 2022).

For **cross-lingual** settings, aligning over diverse language data helps improve low-resource performance without considerably degrading high-resource capabilities (Imani et al., 2023; Blevins and Zettlemoyer, 2022). Parallel corpora allow further gains via translation-based or other cross-lingual masking strategies (Wang et al., 2023a; Conneau and Lample, 2019). However, as the number of languages increases, the curse of dimensionality can dilute representational capacity, making language-specific adaptation or modular training strategies increasingly important (Alabi et al., 2022; Conneau et al., 2020). For **domain adaptation**, CPT on modest amounts of high-quality in-domain text enables models to better capture domain-specific terminology and discourse patterns (Gururangan et al., 2020; Bai et al., 2021). This approach typically outperforms supervised FT alone and is more effective than mixing domain-specific with general data and training from scratch (Türkmen et al., 2023). Targeting specific vocabulary in the replaced or next token prediction objectives can further enhance performance (Lu et al., 2023). In low-resource scenarios, proxy corpora or mixed-domain CPT can help, but the relevance of the data often outweighs its quantity (Mahapatra et al., 2022; Jantscher et al., 2023). In contrast to full encoder PT on billions or trillions of tokens, CPT with as few as 100K tokens can be beneficial, particularly for the newly introduced weights in PEFT methods (Jukić and Snajder, 2023; Gnehm et al., 2022).

Catastrophic forgetting remains a risk, especially for longer training runs with decoders (Kalajdzievski, 2024), requiring regularization techniques (§ 5.4) combined with PEFT methods

(§ 3) to preserve prior knowledge while adapting selectively (Liu et al., 2022; Jukić and Snajder, 2023). In contrast, encoder models benefit even from longer training on small in-domain corpora (Gnehm et al., 2022). Scaling laws for decoder model PT suggest equal scaling of PEFT parameters and data size (Kaplan et al., 2020; Hoffmann et al., 2022). Crucially, careful attention to (inconspicuous) data contamination is essential, especially for public datasets, as the overlap between PT and evaluation sets can artificially inflate results (Ruder et al., 2023).

4.2 Fine-tuning Encoders and Decoders

While encoder and decoder architectures differ in their PT objectives (masked vs. autoregressive language modeling), the challenge of adapting them with limited data shares many similarities. Both require strategies that enhance generalization, mitigate overfitting, and leverage inductive biases from PT. However, some methods naturally align more with one model type: Encoder-oriented techniques often operate in the representation space, while decoder-oriented ones emphasize task diversity and instruction-following. This section surveys approaches for both, highlighting where they converge and where their applicability diverges.

Pattern-exploiting Training. PET reformulates classification as cloze-style language modeling for encoder models, aligning PT and FT objectives (Schick and Schütze, 2021a,b). It relies on handcrafted patterns and verbalizers to map inputs and labels to token sequences, making it effective for few-shot tasks in low-resource domains and languages (Lu et al., 2023; Ullah et al., 2023; Song et al., 2023; Qi et al., 2022). Despite its success, inference often involves costly autoregressive decoding. Recent extensions mitigate this by introducing prototypical decoding (Karimi Mahabadi et al., 2022), label embedding learning, and evolutionary verbalizer optimization (Ling et al., 2023). While PET originated in the encoder setting, its principles (of using prompts to bridge PT and downstream objectives) have influenced instruction-tuning methods for decoders.

Meta-learning. Another line of encoder-focused research improves generalization by learning inductive biases that facilitate rapid adaptation to new discriminative tasks. Metric-based methods encode class-specific prototypes in an embedding space for non-parametric inference (Snell et al., 2017; Wen et al., 2021), while contrastive and triplet-based strategies further refine the latent space (Wu et al., 2024; Pauli et al., 2023). Optimization-based methods, such as model-agnostic meta-learning, learn initialization parameters in a two-level optimization that can quickly adapt to new tasks with few updates (Finn et al., 2017; Bansal et al., 2020; Li et al., 2022; Huang et al., 2023b; Chien et al., 2023). Although less common today, meta-learning's emphasis on representation regularization and rapid adaptation continues to influence newer methods that operate in the representation space of both encoder and decoder models.

Intermediate Fine-tuning. Encoders also benefit from intermediate FT (or task-adaptive PT), where models are fine-tuned on related, label-rich tasks before the target task. This improves transfer by injecting domain-relevant inductive signals (Pruksachatkun et al., 2020; Vu et al., 2020; Phang et al., 2019). A closely related idea in decoder models is multi-task instruction tuning, where the model is exposed to a wide variety of datasets and instructions, scaling the benefits of intermediate supervision to heterogeneous task distributions. This distinction reflects the architectures: encoders profit from staged specialization, whereas decoders gain generalization by broad coverage of instructions.

Multi-task Instruction Tuning. For decoder models, multi-task instruction tuning is the dominant paradigm for generalization. By fine-tuning on diverse datasets and instruction styles, models acquire robust zero- and few-shot instruction-following abilities (Wei et al., 2021; Aribandi et al., 2021; Sanh et al., 2022). Compared to encoder-style intermediate fine-tuning, this approach is broader in scope, exposing decoders to heterogeneous task mixtures that yield checkpoints with faster convergence and stronger generalization to unseen tasks (Longpre et al., 2023; Liu et al., 2022). However, task selection and mixing require care, as diversity can improve generalization but also cause negative task transfer (Wang et al., 2023d; Xia et al., 2024).

Mixture of Experts. To mitigate such trade-offs, mixture-of-experts (MoE) architectures have

emerged as a modular alternative for decoders. Inputs are routed to specialized subnetworks based on input semantics (Shazeer et al., 2017; Fedus et al., 2022), via retrieval (Jang et al., 2023) or dynamic gating mechanisms (Ostapenko et al., 2024; Ponti et al., 2022), improving modularity and continual learning. MoE systems often outperform monolithic instruction-tuned baselines in zero- and few-shot generalization, especially for compositional reasoning (Muqeeth et al., 2024). Recent large-scale decoders combine instruction tuning with massive MoE setups to reduce negative transfer while improving efficiency with fewer active parameters (OpenAI et al., 2025; Meta AI, 2025).

Effective Instruction Design. Decoder performance is also sensitive to instruction formatting. Generalization improves with prompt diversity, sequence packing, and careful task mixture balancing (Sanh et al., 2022; Iyer et al., 2023). Mixed prompting setups (zero-shot, few-shot, chain-of-thought) further enhance robustness (Longpre et al., 2023). Inverting input-output roles and increasing lexical diversity acquire task-agnostic learning abilities (Min et al., 2022). These strategies not only equip models with robust in-context learning and reasoning abilities (critical for low-shot generalization; Wei et al., 2022; Liu et al., 2023d), but also accelerate convergence in subsequent specialized FT (Longpre et al., 2023).

To Mask or Not to Mask? A subtle design choice in decoder fine-tuning is whether to apply loss only on output tokens or also on prompt tokens. While the dominant practice masks loss on instruction tokens (Grattafiori et al., 2024; Ouyang et al., 2022; Wei et al., 2021), recent findings show that including prompt tokens as regularization improves generalization when instructions are long and data is limited ($<$15k examples; Shi et al., 2024; Huerta-Enochian and Ko, 2024). Other forms of regularization, such as KL constraints against a base model, may hurt instruction-following, while noise injection shows mixed results (Jain et al., 2024).

5 Fine-tuning for Specialization

Fine-tuning for specialization focuses on adapting models to domain- or language-specific tasks that fall outside general capabilities (Table 1).

5.1 Embedding Learning

Since fixed vocabularies limit generalization to new domains and languages, recent work retrains or augments token embeddings (often with frozen Transformer body) to better match target data distributions (Hung et al., 2023; Artetxe et al., 2020). This can bridge vocabulary gaps efficiently, particularly in low-resource or cross-lingual settings. Subword-level tokenization remains standard for balancing efficiency and expressiveness (Kudo and Richardson, 2018), but it can fragment rare or domain-specific words. Entropy-based methods identify such tokens and selectively expand the vocabulary with new embeddings (Nag et al., 2023). Alternative strategies map embeddings from high-resource to low-resource languages via seed dictionaries or aligned corpora (Mikolov et al., 2013; Minixhofer et al., 2022), or learn shared embedding spaces through alignment objectives (Cao et al., 2019; Saadi et al., 2022).

5.2 Contrastive and Adversarial Learning

Contrastive and adversarial objectives enhance adaptation by improving representation alignment across domains and languages and are compatible with PEFT methods (§ 3).

Contrastive Learning. This pulls semantically similar pairs closer while pushing apart unrelated ones, improving both sentence- and word-level alignment in cross-lingual transfer (Chi et al., 2021a,b; Gaschi et al., 2023; Hu et al., 2023a). It can be used on parallel corpora or via data augmentation for consistency, and integrates well with PEFT methods like adapters (Liu et al., 2023c,b; Ahmat et al., 2023). In downstream tasks, contrastive objectives boost performance for classification and similarity tasks using class descriptions, anchor prompts, or QA reformulations (Gao et al., 2021; Pauli et al., 2023; Chen et al., 2023c).

Adversarial Learning. This improves generalization by training a discriminator that the model must deceive, encouraging domain- or language-invariant representations (Du et al., 2020; Grießhaber et al., 2020; Huang et al., 2023c). It is especially effective in the absence of parallel data and complements PEFT via language-specific adapters or domain-aware prefix tuning (Ngo Trung et al., 2021; Zhao et al., 2022).

5.3 Limited Supervision

In low-resource scenarios, semi-supervised, unsupervised, and active learning methods can successfully leverage unlabeled data to boost model generalization and robustness.

Semi-supervised Learning. This uses unlabeled data to augment small labeled sets. Common approaches include self-training with pseudo-labels filtered by model confidence or entropy (Schick and Schütze, 2021a; Wang et al., 2023b; Chen et al., 2020), and consistency regularization to encourage stable predictions under perturbations (Sohn et al., 2020; Xie et al., 2020). Co-training variants train multiple views of the data to improve robustness (Clark et al., 2018). These methods are particularly effective in low-resource adaptation, where pre-trained models provide valuable supervision signals to extend the limited annotated data.

Unsupervised Methods. These often rely on consistency conditions such as input perturbation or cycle-consistency across modalities or domains (Zhu et al., 2017). These methods enable learning bidirectional relationships without parallel data, useful for cross-domain, cross-lingual, or style transfer settings (Lample et al., 2018; Ren et al., 2019; Karisani, 2022). However, they typically require large unlabeled corpora and are most effective when combined with supervised signals.

Active Learning (AL). AL reduces annotation cost by selecting the most informative examples for labeling, typically using uncertainty (e.g., confidence, entropy, perplexity) or diversity (e.g., gradient similarity) criteria (Muradoglu and Hulden, 2022; Yuan et al., 2020; François and Gay, 2023; Karamcheti et al., 2021). Iterative sampling and model re-initialization improve stability (Lemmens and Daelemans, 2023). AL reduces the amount of labeled data needed for classification task sometimes even to 30% of the original size; can be beneficial above 1k samples, below any aditional labeled data works (Lemmens and Daelemans, 2023). It integrates well with PEFT methods like adapters and UniPELT, boosting sample efficiency in specialized tasks (Jukić and Snajder, 2023). The focus of AL on selective data efficiency aligns with findings that data quality often outweighs quantity in low-resource adaptation (see § 5.4 and § 6.2).

5.4 Instruction Tuning for Specialization

Instruction tuning specializes decoder models by aligning them with user intent and leveraging structured task descriptions.[1] Beyond multi-task setups aimed at generalization (§ 4.2), clear, specific, and contextually grounded instructions consistently improve task performance, even under severe data constraints (Wang et al., 2023c; Fleming et al., 2024; Zhang et al., 2023b). In contrast to the diverse **instruction design** required by multi-task models (§ 4.2), instruction format variety only brings marginal gain in task-specific performance, and can even be counterproductive in some cases (Chen et al., 2023a). The lack of consistency between train and evaluation prompts can lead to performance drops, larger models being more robust to unseen formats (Gupta et al., 2023).

Specialization vs. Generalization. Just as large-scale multi-task instruction tuning can hurt single-task performance, specializing too narrowly can reduce generalization (Wang et al., 2023d; Zhang et al., 2023a; Alabi et al., 2022). The best way to address this is by fine-tuning instruction-tuned models on the downstream task, which drastically improves sample efficiency, converges faster, and leads to overall better performance (Gupta et al., 2023; Longpre et al., 2023). Adding general instruction data during SFT and applying regularization (§ 5.4) help maintain broader capabilities.

Scaling Models and Data. For multilingual closed-generation tasks (like translation or summarization), Zhang et al. (2023a) find that SFT performance scales with data according to a power law with diminishing returns after an elbow point (around 10k–40k samples for PEFT), depending on the task. SFT data has a larger impact on full FT than on PEFT, but full FT becomes justified after 100k to 1M samples. Surprisingly, model size has a larger impact than FT or PT data size and scaling PEFT parameters is virtually ineffective (Zhang et al., 2023a). The rationalization is that PEFT relies heavily on pre-trained knowledge, making model capacity more critical, especially for in-distribution tasks. For more out-of-distribution or highly specific tasks,

[1] We use **SFT** to denote supervised fine-tuning, with or without instructions.

Method	Compute Characteristics	Data Requirements	Pros / Trade-offs
Alignment with RL	High GPU demand; long training time; requires reward model and PPO	Ranked responses + reward model (100k+ samples)	Strong preference modeling, but sensitive to hyperparameters; expensive and unstable
Direct preference alignment	Moderate compute; fast convergence; no reward model; works with PEFT or full finetuning	Ranked or binary preference data (10k+)	Simple pipeline; good performance; KTO uses weak binary signals; JPO compare instruction-response pairs; risk of verbosity exploitation
Reference-free alignment	Low compute; no reference model; efficient for PEFT; single-stage	Preference data + SFT data regularization (<10k)	Very efficient; no KL regularization needed; SimPO strong but lacks theory; vulnerable to reward hacking without regularization
Alignment during SFT	Low GPU cost per step; single-stage (multiple iterations for SPIN)	Curated SFT data; optionally use rejection sampling	No preference data needed; limited by quality of SFT data; SPIN highly sample efficient

Table 2: Taxonomy of preference alignment strategies with practical characteristics.

however, scaling fine-tuning data becomes more important, particularly in the early training phase.

Identifying Influential Data. This is key to developing specialized capabilities efficiently. Targeted instruction tuning can match full-data performance with as few as 10–15k samples, highlighting the value of high-quality supervision (Xia et al., 2024; Chen et al., 2023a). These approaches rely on statistical properties, similarity metrics, or gradient-based signals (Xie et al., 2023; Zhang et al., 2018; Xia et al., 2024), and show that a core subset remains effective across tasks and model scales. Beyond SFT, similar principles have been applied to filter CPT data based on task relevance (Xie et al., 2023; Mahapatra et al., 2022), offering practical tools for denoising, data selection, or active augmentation (§ 5.3).

Catastrophic Forgetting Mitigation. To prevent overfitting and preserve pre-trained knowledge during SFT with limited data, regularization techniques are vital. Early methods like layer-wise learning rate decay (LLRD), weight decay, and Mixout remain widely used, especially in encoder models (Howard and Ruder, 2018; Wiese et al., 2017; Lee et al., 2019). For instruction tuning tasks with short outputs (e.g., classification, QA), applying a small loss weight (0.01–0.1) to the prompt tokens helps to avoid overfitting (Huerta-Enochian and Ko, 2024). KL divergence is commonly used in LLM alignment to constrain outputs, but less so in standard SFT (Ouyang et al., 2022). Regardless of model size, learning rate warmup and small batch sizes consistently improve training stability with limited data (Buonocore et al., 2023; Xu et al., 2021a). Dropout and gradient clipping are also standard practices (Wolf et al., 2020; von Werra et al., 2020).

Hyperparameters for Low-resource SFT. Despite task and model differences, several practical defaults consistently yield strong performance in low-resource SFT. Learning rates between 5e-6 and 5e-5 are typical, with 2e-5 proving effective across popular models like Alpaca and Llama (Pareja et al., 2024; Jin et al., 2023; Grattafiori et al., 2024). When tuning, scaling the learning rate by a factor of 3 is a simple yet effective heuristic. Small batch sizes (1–8 per GPU) with gradient accumulation are favored for memory efficiency and better generalization (Ding et al., 2023b; Xu et al., 2021a). While 2–3 epochs often suffice, training longer (up to 20–25 epochs) can help in extreme low-data regimes, provided early stopping is used to prevent overfitting (Muennighoff et al., 2023a; Liu et al., 2024a). Other stable choices include AdamW, cosine decay learning rate schedule with warmup ratios of 3–5%, and gradient clipping set to 1. LoRA fine-tuning commonly uses rank and alpha values of 16 with 0.05 dropout (von Werra et al., 2020).

6 Preference Alignment

While fine-tuning equips PLMs with task- and domain-specific knowledge, it often falls short of aligning outputs with nuanced human preferences, especially when data is limited (Ganguli et al., 2022; Bai et al., 2022a). PA (Table 2) addresses this gap by training models with ranked or scored

	Method	Key idea	Training Objective	Hyperparameters
RL § 6	RM+PPO (Ziegler et al., 2020)	Learn a reward model and use it to guide output generation via RL	$-r_\phi(x,y) + \beta \log \frac{\pi_\theta(y\|x)}{\pi_{\text{ref}}(y\|x)}$, where $r_\phi(x,y) = \mathbb{E}_{(x,y_w,y_l)\sim\mathcal{D}}\left[\log \sigma\left(r_\phi(x,y_w) - r_\phi(x,y_l)\right)\right]$	$\beta \in \{0.001, 0.01, 0.05\}$
	d-RLAIF (Lee et al., 2024)	Use LLM-generated preference scores as a reward to guide PPO training	$-s(x,y) + \beta \log \frac{\pi_\theta(y\|x)}{\pi_{\text{ref}}(y\|x)}$, where $s(x,y)$ is the normalized score produced by an off-the-shelf LLM	$\beta \in \{0.001, 0.01, 0.05\}$
Direct § 6	DPO (Rafailov et al., 2023)	Prefer better responses via a logistic loss on reference-adjusted likelihoods	$-\log \sigma\left(\beta \log \frac{\pi_\theta(y_w\|x)}{\pi_{\text{ref}}(y_w\|x)} - \beta \log \frac{\pi_\theta(y_l\|x)}{\pi_{\text{ref}}(y_l\|x)}\right)$	$\beta \in \{0.01, 0.025, 0.05, 0.1\}$
	IPO (Azar et al., 2024)	Regularize DPO to avoid overfitting to extreme preference scores	$\left(\log \frac{\pi_\theta(y_w\|x)}{\pi_{\text{ref}}(y_w\|x)} - \log \frac{\pi_\theta(y_l\|x)}{\pi_{\text{ref}}(y_l\|x)} - \frac{1}{2\tau}\right)^2$	$\tau \in \{0.01, 0.025, 0.0375, 0.5\}$
	R-DPO (Park et al., 2024)	Penalize verbosity that can mislead preference learning	$-\log \sigma\left(\beta \log \frac{\pi_\theta(y_w\|x)}{\pi_{\text{ref}}(y_w\|x)} - \beta \log \frac{\pi_\theta(y_l\|x)}{\pi_{\text{ref}}(y_l\|x)} + (\alpha\|y_w\| - \alpha\|y_l\|)\right)$	$\alpha \in \{0.05, 0.1, 0.5, 1.0\}$, $\beta \in \{0.01, 0.05, 0.1\}$
	KTO (Ethayarajh et al., 2024)	Use unpaired binary feedback to optimize for coherence with reference distribution	$-\lambda_w \sigma\left(\beta \log \frac{\pi_\theta(y_w\|x)}{\pi_{\text{ref}}(y_w\|x)} - z_{\text{ref}}\right) + \lambda_l \sigma\left(z_{\text{ref}} - \beta \log \frac{\pi_\theta(y_l\|x)}{\pi_{\text{ref}}(y_l\|x)}\right)$, where $z_{\text{ref}} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\beta \text{KL}(\pi_\theta(y\|x)\|\pi_{\text{ref}}(y\|x))]$	$\lambda_l = \lambda_w = 1.0$, $\beta \in \{0.01, 0.05, 0.1\}$
	JPO (Bansal et al., 2025)	Generalize preference comparison to arbitrary input pairs	$-\log \sigma\left(\beta \log \frac{\pi_\theta(y_w\|x_w)}{\pi_{\text{ref}}(y_w\|x_w)} - \beta \log \frac{\pi_\theta(y_l\|x_l)}{\pi_{\text{ref}}(y_l\|x_l)}\right)$	$\beta = 0.01$
Reference-free § 6	RRHF (Yuan et al., 2023)	Simple regularised ranking-based alignment with length normalization	$\max\left(0, -\frac{1}{\|y_w\|}\log \pi_\theta(y_w\|x) + \frac{1}{\|y_l\|}\log \pi_\theta(y_l\|x)\right) - \lambda \log \pi_\theta(y_w\|x)$	$\lambda \in \{0.1, 0.5, 1.0, 10.0\}$
	SLiC-HF (Zhao et al., 2023b)	Calibrate sequence likelihood with a regularized, ranked loss	$\max(0, \delta - \log \pi_\theta(y_w\|x) + \log \pi_\theta(y_l\|x)) - \lambda \log \pi_\theta(y_w\|x)$	$\delta = 1, \lambda \in \{0.1, 0.5, 1.0, 10.0\}$
	CPO (Xu et al., 2024)	Removing the reference from the DPO loss leads to an efficient upper bound	$-\log \sigma\left(\beta \log \pi_\theta(y_w\|x) - \beta \log \pi_\theta(y_l\|x)\right) - \log \pi_\theta(y_w\|x)$	$\lambda = 1.0$, $\beta \in \{0.01, 0.05, 0.1\}$
	ORPO (Hong et al., 2024)	Discourage undesired behaviors during SFT with an odds-ratio loss.	$-\log p_\theta(y_w\|x) - \lambda \log\left(\frac{p_\theta(y_w\|x)}{1-p_\theta(y_w\|x)} - \log \frac{p_\theta(y_l\|x)}{1-p_\theta(y_l\|x)}\right)$, where $p_\theta(y\|x) = \exp\left(\frac{1}{\|y\|}\log \pi_\theta(y\|x)\right)$	$\lambda \in \{0.1, 0.5, 1.0, 2.0\}$
	SimPO (Meng et al., 2024)	Align reward function with the generation metric via length-aware, reference-free reward margin	$-\log \sigma\left(\frac{\beta}{\|y_w\|}\log \pi_\theta(y_w\|x) - \frac{\beta}{\|y_l\|}\log \pi_\theta(y_l\|x) - \gamma\right)$	$\beta \in \{2.0, 2.5\}$, $\gamma \in \{0.3, 0.5, 1.0, 1.2, 1.4, 1.6\}$
SFT § 6	CoH (Liu et al., 2023a)	Teach preference via causal modeling over concatenated bad → good examples	$-\log \pi_\theta\left(y_{\text{bad}},\ y_{\text{good}} \mid x\right)$, where input is $(x,\ y_{\text{bad}},\ y_{\text{good}})$	–
	SPIN (Chen et al., 2024b)	Iteratively fine-tune using self-play to match human-like responses	$-\log \sigma\left(\lambda \log \frac{\pi_\theta(y_w\|x)}{\pi_{\theta_t}(y_w\|x)} - \lambda \log \frac{\pi_\theta(y\|x)}{\pi_{\theta_t}(y\|x)}\right)$	$\lambda \in \{5.0$ (last iter.), 0.1 (else)$\}$

Table 3: Objectives and design motivations of various preference alignment strategies, with representative hyperparameters. See § 6 dor detailed discussions.

feedback to better capture task-specific objectives like structure, helpfulness, truthfulness, or biases (Ziegler et al., 2020; Christiano et al., 2017; Ramamurthy et al., 2022). This step usually follows SFT, once the model has already adapted to the task and domain. Without this prior adaptation, models tend to struggle with preference learning and exhibit higher rates of hallucination (Tunstall et al., 2024; Ethayarajh et al., 2024; Hong et al., 2024). Refer to Table 3 for a concise overview of preference alignment objectives.

Alignment with RL. Reinforcement Learning from Human Feedback (RLHF) aligns model behavior by training a reward model to score outputs based on human preferences, followed by optimizing the model using RL algorithms like PPO (Ouyang et al., 2022; Ziegler et al., 2020; Schulman et al., 2017). While effective, this process is resource-intensive (requiring large-scale human annotations, careful hyperparameter tuning, and substantial compute), making it inaccessible for most organizations (Touvron et al.,

2023; Rafailov et al., 2023; Gao et al., 2023). Variants like Reinforcement Learning from AI Feedback (RLAIF) replace human-labeled rewards with scores from an LLM during training, reducing cost while maintaining comparable performance (Lee et al., 2024). However, RLAIF can be less coherent than RLHF, and both methods often produce longer outputs than SFT models, with occasional hallucinations. Despite RLHF's strengths, its instability and cost have spurred interest in more stable and efficient PA alternatives.

Direct Preference Alignment. Recent methods simplify PA by eliminating the need for separate reward models and RL. Direct Preference Optimization (DPO; Rafailov et al., 2023) reframes alignment as a binary classification problem over preference pairs, avoiding LM sampling, reward model training, and heavy hyperparameter tuning. Several extensions build on DPO to improve generalization or efficiency. Identity Preference Optimization (IPO; Azar et al., 2024) introduces regularization to reduce overfitting. Kahneman-Tversky Optimization (KTO; Ethayarajh et al., 2024) removes the need for paired preferences, relying only on binary feedback (good/bad). Length-regularized DPO (R-DPO; Park et al., 2024) penalizes verbosity that may superficially align with human preferences. Joint Preference Optimization (JPO; Bansal et al., 2025) generalizes preference comparison beyond matching prompt inputs, capturing richer human judgments. These approaches offer practical, scalable alternatives to RL-based methods while maintaining strong alignment performance.

Reference-free Alignment. These methods remove the need for a reference model by using alternative regularization strategies to keep outputs close to the SFT distribution and avoid degradation. This significantly reduces memory and compute costs, making these methods attractive for resource-constrained settings. Contrastive preference optimization (CPO; Xu et al., 2024) approximates DPO by replacing the reference model with a uniform prior, using standard NLL loss for regularization. ORPO (Hong et al., 2024) merges SFT and alignment into a single step using an odds ratio loss. ORPO avoids the need for a reference model and separate FT step, offering greater efficiency than DPO or RLHF. SimPO (Meng et al., 2024) addresses misalignment between preference signals and generation behavior by using a normalized, reference-free reward based on likelihood and output length. A reward margin further sharpens the model's ability to distinguish between preferred and rejected responses. These methods offer lightweight alternatives to DPO and RLHF, especially suited for practical deployments in constrained environments.

Alignment During SFT. Another direction integrates PA directly into the SFT stage, removing the need for additional alignment steps. Careful curation of SFT datasets, such as filtering for specific structure or style, can yield strong alignment performance on its own, without dedicated preference data (Zhou et al., 2023; Li et al., 2023; Dong et al., 2023; Yuan et al., 2023). Chain of Hindsight (Liu et al., 2023a) presents both a good and bad response during training and teaches the model to prefer the good one at inference time, embedding preference signals without explicit reward modeling. SPIN (Chen et al., 2024b) uses self-play to iteratively fine-tune a model to distinguish human-written and self-generated responses, gradually aligning its output distribution to human-like behavior. These methods rely on high-quality SFT data and can serve as effective precursors to alignment techniques like DPO, emphasizing that data quality (and not quantity) is often the limiting factor (§ 6.2), and that well-filtered SFT can be a lightweight, effective form of alignment.

6.1 Choosing an Alignment Strategy

PPO offers strong alignment performance but is resource-intensive and unstable, requiring large datasets and extensive hyperparameter tuning (Ouyang et al., 2022; Rafailov et al., 2023). This makes it impractical in many cases. Nevertheless, reward modeling itself has been shown to be relatively robust to moderate levels of label noise (up to 30%; Shen et al., 2024).

DPO and its variants eliminate the reward model and are more stable. IPO, especially, is less prone to reward hacking, achieving lower KL divergence under the same constraints (Rafailov et al., 2024). DPO scales better than PPO to models of size up to 405B, but might require regularization (e.g., additional negative-log-likelihood (NLL) term with a coefficient of 0.2; Grattafiori et al., 2024). Contrary to reward modeling, DPO

variants can be susceptible to label noise, as rankings are rarely flipped during alignment (Chen et al., 2024a). KTO is more sample efficient by learning from weaker binary signals and outperforms DPO in noisy or imbalanced datasets, especially in the 1B–30B parameter range (Ethayarajh et al., 2024). JPO improves data efficiency by directly comparing instruction-response pairs, although a comparison might be ambiguous. In general, masking out formatting tokens is recommended in objectives of contrasting nature to avoid conflicting signals (Grattafiori et al., 2024).

Reference-free methods further reduce compute requirements by dispensing with reference models, and can be particularly effective in low-resource settings. For instance, CPO pushes translation performance beyond DPO using just 22k samples (Xu et al., 2024). ORPO is on par with DPO, but produces more consistent and diverse outputs (Hong et al., 2024). SimPO delivers superior performance across multiple benchmarks, although the underlying reasons for its effectiveness lack theoretical justification (Meng et al., 2024). Despite their efficiency, all reference-free approaches are more susceptible to reward hacking, especially when fine-tuning the full model (Chen et al., 2024a).

Combining alignment objectives with SFT loss improves stability and performance for reasoning or domain-specific tasks (Xu et al., 2024; Meng et al., 2024). Integrating PA into SFT presents a lightweight alternative that requires only high-quality instruction data (Zhou et al., 2023). SPIN, in particular, achieves DPO-level performance with less data, however, it requires multiple training iterations, trading memory efficiency for time (Chen et al., 2024b).

In practice, DPO and its variants remain strong defaults for general use. Reference-free methods are ideal for compute-constrained settings but demand careful regularization. Alignment during SFT is attractive when preference data is unavailable and curation quality is high. Table 2 summarizes trade-offs between alignment techniques, while Table 3 contrasts their training objectives.

6.2 Preference Data Criteria

Format and Availability. Preference datasets typically consist of chosen/rejected pairs derived from instruction-response prompts (Ouyang et al., 2022; Rafailov et al., 2023; Meng et al., 2024). Some approaches extend this to comparisons across prompts (JPO) or simple binary signals on the output (KTO), but the overall availability remains limited, with most datasets ranging from thousands to 1M samples and predominantly in English (Argilla, 2024; Liu, 2025).

Synthetic Feedback as a Scalable Solution. Collecting high-quality human preference data is resource-intensive. A widely adopted alternative is using synthetic feedback from strong LLMs guided by curated principles (Bai et al., 2022b). This enables broader scalability while maintaining quality when human oversight is present. Methods like SPIN go further by generating negative examples automatically via self-play, relying solely on an initial set of desirable outputs (Chen et al., 2024b). KTO also facilitates feedback collection by supporting imbalanced binary labels instead of strict pairwise comparisons (Ethayarajh et al., 2024).

When feedback is derived from strong teacher models (e.g., GPT-4 or Claude), interestingly, SFT has been observed to rival preference-based alignment methods like DPO or RL (Sharma et al., 2024). This can be partly due to insufficient diversity or a weak student model not exceeding the quality of the synthetic data during exploration. Diminishing returns beyond a few thousand samples suggest the importance of selection and coverage rather than raw volume.

Data Quality over Quantity. Multiple studies consistently show that a small number of high-quality, diverse examples outperforms larger, noisier datasets (Zhou et al., 2023; Shen et al., 2024). For instance, SPIN matches its original 50k-sample performance using just 1.8k well-curated examples, and ORPO remains effective at 141B scale with only 7k samples (Argilla and MantisNLP, 2024). Similar trends hold across alignment methods: as few as 1k–10k diverse, well-labeled examples can outperform tens or hundreds of thousands of lower-quality instances (Rafailov et al., 2023; Zhou et al., 2023; Shen et al., 2024). For certain desired behaviors, like adhering to output structure specifications or coherent multi-turn conversations, a handful of high-quality examples can remarkably boost alignment (Zhou et al., 2023). These results underscore the value of filtering and targeted data collection over naive scaling.

Both human-labeled and synthetic preference datasets frequently contain noise, contradictions, or intransitive preferences (Ethayarajh et al., 2024). These issues can undermine alignment objectives unless mitigated by clear annotation guidelines and consistent labeling practices. Additionally, prompting strategies like CoT can help improve the reliability of synthetic data, though techniques like few-shot prompting yield mixed results (Lee et al., 2024).

6.3 Scaling Dynamics for Model and Data

PA performance generally follows a logarithmic scaling trend with respect to both model size and dataset size (Gao et al., 2023). For RLHF, increasing the size of the reward or policy model yields diminishing returns, and more data leads to better generalization but with progressively smaller gains. Notably, larger models do not overfit faster nor diverge more from the initial policy, contradicting earlier intuitions. A minimum dataset size threshold is required to surpass near-random alignment performance. For example, at least 2k preference pairs are necessary for PPO to move beyond baseline loss, independent of the reward model size. Moreover, data scaling is generally more impactful than reward model scaling: A 2-fold increase in data is better than a 4-fold increase in model size. For instance, a 680M model trained on 16k (32k) samples outperforms a 3B model with only 8k (16k) (see Figure 10 in Gao et al., 2023).

In both RLHF and direct alignment methods, higher KL divergence budgets tend to degrade performance, often even before a full epoch is completed. Models frequently reach optimal performance at partial epochs (e.g., $\sim$23k samples) with lower KL constraints (Rafailov et al., 2024). Smaller models, under the same constraints, achieve higher KL divergence and exhibit reward hacking earlier, while models around 7B parameters or more show significantly improved sample efficiency and stability. With limited capacity or overly strict KL constraints, models tend to over-exploit spurious cues like verbosity. Length regularization alone does not fully prevent such behaviors (Rafailov et al., 2024).

Over-saturating SFT with synthetic completions can lead to marginal gains on alignment benchmarks, as observed with LLaMA-7B performance plateauing beyond 5k examples (Sharma et al., 2024). This suggests that, beyond a point, SFT may be insufficient to further improve alignment, especially when data quality is limited, highlighting the need for more targeted alignment methods.

6.4 Effect on Downstream Task Performance

PA can have mixed effects on downstream task performance, depending on the alignment method, model initialization, and task type. While alignment methods often improve truthfulness, commonsense knowledge and reasoning likely due to overlap with preference data, they may slightly reduce performance on knowledge-intensive tasks (Meng et al., 2024). For instance, mathematical reasoning performance can significantly deteriorate for methods without a regularization SFT loss term, like CPO or ORPO. Tunstall et al. (2024) report that DPO alignment starts to hurt downstream performance after a single epoch.

Preference-aligned models often produce longer outputs than those trained with SFT alone, even when the objective includes length regularization (Park et al., 2024; Meng et al., 2024). However, verbosity shows weak correlation with KL divergence, implying that human preferences also depend on factors like repetition and tone, which may affect downstream performance. Overall, the effects of alignment strategies on downstream outcomes remain underexplored, particularly in relation to model scale, output style, and non-English settings, highlighting a need for more systematic evaluation.

7 Discussion

Encode or Decode? Choosing the appropriate model architecture is critical for optimizing performance across NLP tasks, particularly in resource-constrained or domain-specific settings (Alabi et al., 2020). While large decoder-only models such as GPT-4 or LLaMA can be adapted to any NLP task via prompt-based reformulations (Liu et al., 2023d), they often require substantial computational resources for training and inference. In contrast, encoder-only models like RoBERTa or DeBERTa (Liu et al., 2019; He et al., 2020) are generally more efficient for tasks centered on natural language understanding (NLU), such as classification, named entity recognition, and extractive QA, often outperforming decoder models with far fewer parameters (Benayas et al.,

2024; Qorib et al., 2024; García-Díaz et al., 2023; Schick and Schütze, 2021a). Their bidirectional attention and dense token supervision (i.e., predicting multiple masked tokens per input sequence) during training enhances sample efficiency and makes them especially effective in low-resource scenarios (Ding et al., 2023b; Karimi Mahabadi et al., 2022). For instance, RoBERTa (335M) achieves performance close to T5 (11B) despite having an order of magnitude fewer parameters (Ding et al., 2023b). While large decoder models dominate many mainstream applications, encoder models remain highly effective in domains requiring deep semantic understanding, limited supervision, and efficient inference. Their fixed-length embeddings and fast processing make them well-suited for tasks like retrieval, reranking, and domain-specific classification. Ultimately, architecture choice should be driven by task requirements, the trade-off between understanding and generation needs, and resource constraints. Table 4 summarizes effective encoder-based approaches for low-resource discriminative NLP tasks.

Cross-lingual Transfer. This remains essential for deploying LLMs in multilingual settings, especially low-resource languages. Strong zero-shot generalization has been observed even when some languages were minimally present during PT (Armengol-Estapé et al., 2022). Transfer improves with language similarity (via word order or sub-word overlap; Deshpande et al., 2022), and can further be enhanced using PEFT methods like adapters or prefix-tuning (Wu et al., 2023; Tu et al., 2024). Cross-lingual prompting offers a training-free strategy for low-resource languages, with English prompts yielding more consistent results in few-shot settings (Lin et al., 2022; Huang et al., 2023a). Prompting the model to think in English before responding in the target language further improves task performance via CoT reasoning (Qin et al., 2023). Translating training data generally outperforms direct transfer or translating test data in domain-specific tasks (Gaschi et al., 2023). Finally, SFT using multilingual prompts (not just English) leads to better cross-lingual generalization (Muennighoff et al., 2023b).

Model Merging. This has emerged as a practical and effective strategy to enhance LLM capabilities without requiring additional training or ensembling overhead. By operating directly in weight space, merging enables the combination of models trained under different configurations, such as various data, hyperparameters, or training objectives (e.g., CPT, SFT, DPO/ORPO) into a single, stronger model (Lu et al., 2025; Goddard et al., 2024). Techniques like Model Soup (Wortsman et al., 2022) and SLERP (Shoemake, 1985) allow smooth interpolation between models, often leading to better downstream performance than any individual component. Notably, model merging contributed to top-ranked models on the Open LLM Leaderboard, being frequently adopted for the state of the art (Grattafiori et al., 2024). It has also proven effective for low-resource languages, where merging models with complementary capabilities outperforms conventional CPT followed by SFT (Tao et al., 2024). However, such benefits appear limited in smaller models (e.g., $<$2B parameters), suggesting scaling is crucial for emergent behaviors (Lu et al., 2025).

8 Conclusion

Fine-tuning language models under data-scarce conditions remains a central challenge for practitioners and researchers aiming to build high-performing systems without the extensive resources typically required for large-scale training. This survey presents a structured and pragmatic overview of the entire post-pretraining development pipeline, reviewing and organizing methods across parameter-efficient fine-tuning, domain and cross lingual generalization, specialization, and preference alignment. We highlight promising use cases, trade-offs, and best practices grounded in current empirical evidence. Our findings underscore that larger models are consistently more sample-efficient. Further, data quality remains more critical than quantity, supporting the view that most capabilities are learned during pre-training, with fine-tuning and alignment primarily shaping model behavior. We emphasize that despite the growing dominance of decoder-based models, encoder models remain competitive for discriminative tasks, and hybrid architectures provide promising avenues (e.g., RAG). In preference alignment, the lack of correlation between alignment loss and task performance (Rafailov et al., 2024) points to a deeper calibration issue that requires better metrics and theoretical grounding. Model merging

can be surprisingly effective, yet the nonlinear interactions underlying this behavior are poorly understood, motivating further research. Finally, we stress the importance of transparent reporting of datasets, training protocols, and evaluation metrics for reproducibility and fair comparison, especially for non-English and domain-specific tasks, where current benchmarks often fall short for assessing real-world applicability.

Acknowledgments

The authors would like to sincerely thank the action editor, Sebastian Ruder, and the anonymous reviewers for their insightful and constructive comments, which significantly helped improve the clarity and practical relevance of this work. This research was supported by the Nemetschek Innovation Foundation and the Bavarian Ministry of Science, Research and Art.

References

Armen Aghajanyan, Sonal Gupta, and Luke Zettlemoyer. 2021. Intrinsic dimensionality explains the effectiveness of language model fine-tuning. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7319–7328, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.568

Ahtamjan Ahmat, Yating Yang, Bo Ma, Rui Dong, Kaiwen Lu, and Lei Wang. 2023. WAD-X: Improving zero-shot cross-lingual transfer via adapter-based word alignment. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(9):224:1–224:23. https://doi.org/10.1145/3610289

Jesujoba Alabi, Kwabena Amponsah-Kaakyire, David Adelani, and Cristina España-Bonet. 2020. Massive vs. curated embeddings for low-resourced languages: The case of Yorùbá and Twi. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2754–2762, Marseille, France. European Language Resources Association.

Jesujoba O. Alabi, David Ifeoluwa Adelani, Marius Mosbach, and Dietrich Klakow. 2022. Adapting pre-trained language models to african languages via multilingual adaptive fine-tuning. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 4336–4349, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Alan Ansell, Edoardo Ponti, Anna Korhonen, and Ivan Vulić. 2022. Composable sparse fine-tuning for cross-lingual transfer. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1778–1796, Dublin, Ireland. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.125

Argilla. 2024. Preference datasets for DPO - an Argilla collection.

Argilla and MantisNLP. 2024. RLHF and Alternatives: Overview.

Vamsi Aribandi, Yi Tay, Tal Schuster, Jinfeng Rao, Huaixiu Steven Zheng, Sanket Vaibhav Mehta, Honglei Zhuang, Vinh Q. Tran, Dara Bahri, Jianmo Ni, Jai Gupta, Kai Hui, Sebastian Ruder, and Donald Metzler. 2021. ExT5: Towards extreme multi-task scaling for transfer learning. In *International Conference on Learning Representations*.

Jordi Armengol-Estapé, Ona de Gibert Bonet, and Maite Melero. 2022. On the multilingual capabilities of very large-scale english language models. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 3056–3068, Marseille, France. European Language Resources Association.

Mikel Artetxe, Sebastian Ruder, and Dani Yogatama. 2020. On the cross-lingual transferability of monolingual representations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4623–4637, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.421

Mohammad Gheshlaghi Azar, Zhaohan Daniel Guo, Bilal Piot, Remi Munos, Mark Rowland, Michal Valko, and Daniele Calandriello. 2024. A general theoretical paradigm to understand

learning from human preferences. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, pages 4447–4455. PMLR.

Fan Bai, Alan Ritter, and Wei Xu. 2021. Pre-train or annotate? Domain adaptation with a constrained budget. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5002–5015, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.409

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson Kernion, Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez, Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt, Neel Nanda, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, Ben Mann, and Jared Kaplan. 2022a. Training a helpful and harmless assistant with reinforcement learning from human feedback. ArXiv:2204.05862v1 [cs].

Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, Carol Chen, Catherine Olsson, Christopher Olah, Danny Hernandez, Dawn Drain, Deep Ganguli, Dustin Li, Eli Tran-Johnson, Ethan Perez, Jamie Kerr, Jared Mueller, Jeffrey Ladish, Joshua Landau, Kamal Ndousse, Kamile Lukosuite, Liane Lovitt, Michael Sellitto, Nelson Elhage, Nicholas Schiefer, Noemi Mercado, Nova DasSarma, Robert Lasenby, Robin Larson, Sam Ringer, Scott Johnston, Shauna Kravec, Sheer El Showk, Stanislav Fort, Tamera Lanham, Timothy Telleen-Lawton, Tom Conerly, Tom Henighan, Tristan Hume, Samuel R. Bowman, Zac Hatfield-Dodds, Ben Mann, Dario Amodei, Nicholas Joseph, Sam McCandlish, Tom Brown, and Jared Kaplan. 2022b. Constitutional AI: Harmlessness from AI feedback. ArXiv:2212.08073v1 [cs].

Hritik Bansal, Ashima Suvarna, Gantavya Bhatt, Nanyun Peng, Kai-Wei Chang, and Aditya Grover. 2025. Comparing bad apples to good oranges: Aligning large language models via joint preference optimization. ArXiv:2404.00530v2 [cs]. https://doi.org/10.18653/v1/2025.findings-acl.39

Trapit Bansal, Rishikesh Jha, Tsendsuren Munkhdalai, and Andrew McCallum. 2020. Self-supervised meta-learning for few-shot natural language classification tasks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 522–534, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.38

Mohammad Bavarian, Heewoo Jun, Nikolas Tezak, John Schulman, Christine McLeavey, Jerry Tworek, and Mark Chen. 2022. Efficient training of language models to fill in the middle. ArXiv:2207.14255v1 [cs].

Elad Ben Zaken, Yoav Goldberg, and Shauli Ravfogel. 2022. BitFit: Simple parameter-efficient fine-tuning for transformer-based masked language-models. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1–9, Dublin, Ireland. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-short.1

Alberto Benayas, Miguel Angel Sicilia, and Marçal Mora-Cantallops. 2024. A comparative analysis of encoder-only and decoder-only models in intent classification and sentiment analysis: Navigating the trade-offs in model size and performance. *Language Resources and Evaluation*. https://doi.org/10.21203/rs.3.rs-3865391/v1

Terra Blevins and Luke Zettlemoyer. 2022. Language contamination helps explains the cross-lingual capabilities of English pretrained models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 3563–3574, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.233

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla

Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. ArXiv: 2005.14165v4 [cs].

Tommaso Mario Buonocore, Claudio Crema, Alberto Redolfi, Riccardo Bellazzi, and Enea Parimbelli. 2023. Localizing in-domain adaptation of transformer-based biomedical language models. *Journal of Biomedical Informatics*, 144:104431. https://doi.org/10.1016/j.jbi.2023.104431, PubMed: 37385327

Steven Cao, Nikita Kitaev, and Dan Klein. 2019. Multilingual alignment of contextual word representations. In *International Conference on Learning Representations*.

Angelica Chen, Sadhika Malladi, Lily H. Zhang, Xinyi Chen, Qiuyi Zhang, Rajesh Ranganath, and Kyunghyun Cho. 2024a. Preference learning algorithms do not learn preference rankings. *Advances in Neural Information Processing Systems*, 37:101928–101968. https://doi.org/10.52202/079017-3234

Hao Chen, Yiming Zhang, Qi Zhang, Hantao Yang, Xiaomeng Hu, Xuetao Ma, Yifan Yanggong, and Junbo Zhao. 2023a. Maybe only 0.5% data is needed: A preliminary exploration of low training data instruction tuning. ArXiv:2305.09246v1 [cs].

Jiaao Chen, Derek Tam, Colin Raffel, Mohit Bansal, and Diyi Yang. 2023b. An empirical survey of data augmentation for limited data learning in NLP. *Transactions of the Association for Computational Linguistics*, 11:191–211. https://doi.org/10.1162/tacl_a_00542

Jiaao Chen, Zichao Yang, and Diyi Yang. 2020. MixText: Linguistically-informed interpolation of hidden space for semi-supervised text classification. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2147–2157. Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.194

Jiaao Chen, Aston Zhang, Xingjian Shi, Mu Li, Alex Smola, and Diyi Yang. 2022a. Parameter-efficient fine-tuning design spaces. In *International Conference on Learning Representations*. https://doi.org/10.1093/bioinformatics/btad496

Peng Chen, Jian Wang, Hongfei Lin, Di Zhao, and Zhihao Yang. 2023c. Few-shot biomedical named entity recognition via knowledge-guided instance generation and prompt contrastive learning. *Bioinformatics*, 39(8):btad496.

Yuxuan Chen, David Harbecke, and Leonhard Hennig. 2022b. Multilingual relation classification via efficient and effective prompting. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 1059–1075, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.69

Zixiang Chen, Yihe Deng, Huizhuo Yuan, Kaixuan Ji, and Quanquan Gu. 2024b. Self-play fine-tuning converts weak language models to strong language models. In *Proceedings of the 41st International Conference on Machine Learning*, pages 6621–6642. PMLR.

Zewen Chi, Li Dong, Furu Wei, Nan Yang, Saksham Singhal, Wenhui Wang, Xia Song, Xian-Ling Mao, Heyan Huang, and Ming Zhou. 2021a. InfoXLM: An information-theoretic framework for cross-lingual language model pre-training. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3576–3588, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.naacl-main.280

Zewen Chi, Li Dong, Bo Zheng, Shaohan Huang, Xian-Ling Mao, Heyan Huang, and Furu Wei. 2021b. Improving pretrained cross-lingual language models via self-labeled word alignment. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3418–3430,

Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.265

Jen-Tzung Chien, Ming-Yen Chen, and Jing-Hao Xue. 2023. Learning meta soft prompt for few-shot language models. In *2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 57–62. https://doi.org/10.1109/APSIPAASC58517.2023.10317500

Paul F. Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Alexandra Chronopoulou, Matthew Peters, Alexander Fraser, and Jesse Dodge. 2023. AdapterSoup: Weight averaging to improve generalization of pretrained language models. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 2054–2063, Dubrovnik, Croatia. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-eacl.153

Kevin Clark, Minh-Thang Luong, Christopher D. Manning, and Quoc Le. 2018. Semi-supervised sequence modeling with cross-view training. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1914–1925, Brussels, Belgium. Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-1217

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised Cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.747

Alexis Conneau and Guillaume Lample. 2019. Cross-lingual language model pretraining. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Claudio Crema, Tommaso Mario Buonocore, Silvia Fostinelli, Enea Parimbelli, Federico Verde, Cira Fundarò, Marina Manera, Matteo Cotta Ramusino, Marco Capelli, Alfredo Costa, Giuliano Binetti, Riccardo Bellazzi, and Alberto Redolfi. 2023. Advancing Italian biomedical information extraction with transformers-based models: Methodological insights and multicenter practical application. *Journal of Biomedical Informatics*, 148:104557. https://doi.org/10.1016/j.jbi.2023.104557, PubMed: 38012982

Ameet Deshpande, Partha Talukdar, and Karthik Narasimhan. 2022. When is BERT multilingual? Isolating crucial ingredients for cross-lingual transfer. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3610–3623, Seattle, United States. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.264

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. QLoRA: Efficient finetuning of quantized LLMs. *Advances in Neural Information Processing Systems*, 36:10088–10115.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1* (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Ning Ding, Xingtai Lv, Qiaosen Wang, Yulin Chen, Bowen Zhou, Zhiyuan Liu, and Maosong Sun. 2023a. Sparse low-rank adaptation of pre-trained language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 4133–4145, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.252

Ning Ding, Yujia Qin, Guang Yang, Fuchao Wei, Zonghan Yang, Yusheng Su, Shengding Hu, Yulin Chen, Chi-Min Chan, Weize Chen, Jing Yi, Weilin Zhao, Xiaozhi Wang, Zhiyuan Liu, Hai-Tao Zheng, Jianfei Chen, Yang Liu, Jie Tang, Juanzi Li, and Maosong Sun. 2023b. Parameter-efficient fine-tuning of large-scale pre-trained language models. *Nature Machine Intelligence*, 5(3):220–235. Nature Publishing Group. https://doi.org/10.1038/s42256-023-00626-4

Jesse Dodge, Gabriel Ilharco, Roy Schwartz, Ali Farhadi, Hannaneh Hajishirzi, and Noah Smith. 2020. Fine-tuning pretrained language models: Weight initializations, data orders, and early stopping. ArXiv:2002.06305v1 [cs].

Hanze Dong, Wei Xiong, Deepanshu Goyal, Yihan Zhang, Winnie Chow, Rui Pan, Shizhe Diao, Jipeng Zhang, KaShun Shum, and Tong Zhang. 2023. RAFT: Reward rAnked FineTuning for generative foundation model alignment. *Transactions on Machine Learning Research*.

Chunning Du, Haifeng Sun, Jingyu Wang, Qi Qi, and Jianxin Liao. 2020. Adversarial and domain-aware bert for cross-domain sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4019–4028, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.370

Ali Edalati, Marzieh Tahaei, Ivan Kobyzev, Vahid Partovi Nia, James J. Clark, and Mehdi Rezagholizadeh. 2022. KronA: Parameter efficient tuning with kronecker adapter. ArXiv:2212.10650v1 [cs].

Liat Ein-Dor, Alon Halfon, Ariel Gera, Eyal Shnarch, Lena Dankin, Leshem Choshen, Marina Danilevsky, Ranit Aharonov, Yoav Katz, and Noam Slonim. 2020. Active learning for BERT: An empirical study. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7949–7962, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.638

Kawin Ethayarajh, Winnie Xu, Niklas Muennighoff, Dan Jurafsky, and Douwe Kiela. 2024. KTO: Model alignment as prospect theoretic optimization. ArXiv:2402.01306v4 [cs].

William Fedus, Barret Zoph, and Noam Shazeer. 2022. Switch Transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23(120):1–39.

Steven Y. Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. 2021. A survey of data augmentation approaches for NLP. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 968–988, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.findings-acl.84

Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 1126–1135. PMLR.

Scott L. Fleming, Alejandro Lozano, William J. Haberkorn, Jenelle A. Jindal, Eduardo Pontes Reis, Rahul Thapa, Louis Blankemeier, Julian Z. Genkins, Ethan Steinberg, Ashwin Nayak, Birju S. Patel, Chia-Chun Chiang, Alison Callahan, Zepeng Huo, Sergios Gatidis, Scott J. Adams, Oluseyi Fayanju, Shreya J. Shah, Thomas Savage, Ethan Goh, Akshay S. Chaudhari, Nima Aghaeepour, Christopher D. Sharp, Michael A. Pfeffer, Percy Liang, Jonathan H. Chen, Keith E. Morse, Emma P. Brunskill, Jason A. Fries, and Nigam H. Shah. 2024. MedAlign: A clinician-generated dataset for instruction following with electronic medical records. In *Proceedings of the AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v38i20.30205, PubMed: 41584261

Matthieu François and Paul Gay. 2023. Active learning with few-shot learning for crisis management. In *Proceedings of the 20th International Conference on Content-based Multimedia Indexing*, CBMI '23, pages 233–237, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3617233.3617278

Deep Ganguli, Liane Lovitt, Jackson Kernion, Amanda Askell, Yuntao Bai, Saurav Kadavath,

Ben Mann, Ethan Perez, Nicholas Schiefer, Kamal Ndousse, Andy Jones, Sam Bowman, Anna Chen, Tom Conerly, Nova DasSarma, Dawn Drain, Nelson Elhage, Sheer El Showk, Stanislav Fort, Zac Hatfield-Dodds, Tom Henighan, Danny Hernandez, Tristan Hume, Josh Jacobson, Scott Johnston, Shauna Kravec, Catherine Olsson, Sam Ringer, Eli Tran-Johnson, Dario Amodei, Tom Brown, Nicholas Joseph, Sam McCandlish, Chris Olah, Jared Kaplan, and Jack Clark. 2022. Red teaming language models to reduce harms: Methods, scaling behaviors, and lessons learned. *CoRR*.

Bofei Gao, Feifan Song, Yibo Miao, Zefan Cai, Zhe Yang, Liang Chen, Helan Hu, Runxin Xu, Qingxiu Dong, Ce Zheng, Shanghaoran Quan, Wen Xiao, Ge Zhang, Daoguang Zan, Keming Lu, Bowen Yu, Dayiheng Liu, Zeyu Cui, Jian Yang, Lei Sha, Houfeng Wang, Zhifang Sui, Peiyi Wang, Tianyu Liu, and Baobao Chang. 2024. Towards a unified view of preference learning for large language models: A survey. ArXiv:2409.02795v5 [cs].

Leo Gao, John Schulman, and Jacob Hilton. 2023. Scaling laws for reward model overoptimization. In *Proceedings of the 40th International Conference on Machine Learning*, pages 10835–10866. PMLR.

Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. SimCSE: Simple contrastive learning of sentence embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6894–6910, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.552

José Antonio García-Díaz, Ronghao Pan, and Rafael Valencia-García. 2023. Leveraging zero and few-shot learning for enhanced model generality in hate speech detection in Spanish and English. *Mathematics*, 11(24):5004. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/math11245004

Felix Gaschi, Patricio Cerda, Parisa Rastin, and Yannick Toussaint. 2023. Exploring the relationship between alignment and cross-lingual transfer in multilingual transformers. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 3020–3042, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-acl.189

Ann-Sophie Gnehm, Eva Bühlmann, and Simon Clematide. 2022. Evaluation of transfer learning and domain adaptation for analyzing german-speaking job advertisements. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 3892–3901, Marseille, France. European Language Resources Association.

Charles Goddard, Shamane Siriwardhana, Malikeh Ehghaghi, Luke Meyers, Vladimir Karpukhin, Brian Benedict, Mark McQuade, and Jacob Solawetz. 2024. Arcee's MergeKit: A toolkit for merging large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 477–485, Miami, Florida, US. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-industry.36

Koustava Goswami, Lukas Lange, Jun Araki, and Heike Adel. 2023. SwitchPrompt: Learning domain-specific gated soft prompts for classification in low-resource domains. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 2689–2695, Dubrovnik, Croatia. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.eacl-main.197

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, Aurelien Rodriguez, Austen Gregerson, Ava Spataru, Baptiste Roziere, Bethany Biron, Binh Tang, Bobbie Chern, Charlotte Caucheteux, Chaya Nayak, Chloe Bi, Chris Marra, Chris McConnell, Christian Keller, Christophe Touret, Chunyang Wu, Corinne Wong, Cristian Canton Ferrer, Cyrus Nikolaidis, Damien Allonsius, Daniel Song, Danielle Pintz, Danny Livshits, Danny Wyatt, David Esiobu, Dhruv Choudhary, Dhruv Mahajan, Diego

Garcia-Olano, Diego Perino, Dieuwke Hupkes, Egor Lakomkin, Ehab AlBadawy, Elina Lobanova, Emily Dinan, Eric Michael Smith, Filip Radenovic, Francisco Guzmán, Frank Zhang, Gabriel Synnaeve, Gabrielle Lee, Georgia Lewis Anderson, Govind Thattai, Graeme Nail, Gregoire Mialon, Guan Pang, Guillem Cucurell, Hailey Nguyen, Hannah Korevaar, Hu Xu, Hugo Touvron, Iliyan Zarov, Imanol Arrieta Ibarra, Isabel Kloumann, Ishan Misra, Ivan Evtimov, Jack Zhang, Jade Copet, Jaewon Lee, Jan Geffert, Jana Vranes, Jason Park, Jay Mahadeokar, Jeet Shah, Jelmer van der Linde, Jennifer Billock, Jenny Hong, Jenya Lee, Jeremy Fu, Jianfeng Chi, Jianyu Huang, Jiawen Liu, Jie Wang, Jiecao Yu, Joanna Bitton, Joe Spisak, Jongsoo Park, Joseph Rocca, Joshua Johnstun, Joshua Saxe, Junteng Jia, Kalyan Vasuden Alwala, Karthik Prasad, Kartikeya Upasani, Kate Plawiak, Ke Li, Kenneth Heafield, Kevin Stone, Khalid El-Arini, Krithika Iyer, Kshitiz Malik, Kuenley Chiu, Kunal Bhalla, Kushal Lakhotia, Lauren Rantala-Yeary, Laurens van der Maaten, Lawrence Chen, Liang Tan, Liz Jenkins, Louis Martin, Lovish Madaan, Lubo Malo, Lukas Blecher, Lukas Landzaat, Luke de Oliveira, Madeline Muzzi, Mahesh Pasupuleti, Mannat Singh, Manohar Paluri, Marcin Kardas, Maria Tsimpoukelli, Mathew Oldham, Mathieu Rita, Maya Pavlova, Melanie Kambadur, Mike Lewis, Min Si, Mitesh Kumar Singh, Mona Hassan, Naman Goyal, Narjes Torabi, Nikolay Bashlykov, Nikolay Bogoychev, Niladri Chatterji, Ning Zhang, Olivier Duchenne, Onur Çelebi, Patrick Alrassy, Pengchuan Zhang, Pengwei Li, Petar Vasic, Peter Weng, Prajjwal Bhargava, Pratik Dubal, Praveen Krishnan, Punit Singh Koura, Puxin Xu, Qing He, Qingxiao Dong, Ragavan Srinivasan, Raj Ganapathy, Ramon Calderer, Ricardo Silveira Cabral, Robert Stojnic, Roberta Raileanu, Rohan Maheswari, Rohit Girdhar, Rohit Patel, Romain Sauvestre, Ronnie Polidoro, Roshan Sumbaly, Ross Taylor, Ruan Silva, Rui Hou, Rui Wang, Saghar Hosseini, Sahana Chennabasappa, Sanjay Singh, Sean Bell, Seohyun Sonia Kim, Sergey Edunov, Shaoliang Nie, Sharan Narang, Sharath Raparthy, Sheng Shen, Shengye Wan, Shruti Bhosale, Shun Zhang, Simon Vandenhende, Soumya Batra, Spencer Whitman, Sten Sootla, Stephane Collot, Suchin Gururangan, Sydney Borodinsky, Tamar Herman, Tara Fowler, Tarek Sheasha, Thomas Georgiou, Thomas Scialom, Tobias Speckbacher, Todor Mihaylov, Tong Xiao, Ujjwal Karn, Vedanuj Goswami, Vibhor Gupta, Vignesh Ramanathan, Viktor Kerkez, Vincent Gonguet, Virginie Do, Vish Vogeti, Vítor Albiero, Vladan Petrovic, Weiwei Chu, Wenhan Xiong, Wenyin Fu, Whitney Meers, Xavier Martinet, Xiaodong Wang, Xiaofang Wang, Xiaoqing Ellen Tan, Xide Xia, Xinfeng Xie, Xuchao Jia, Xuewei Wang, Yaelle Goldschlag, Yashesh Gaur, Yasmine Babaei, Yi Wen, Yiwen Song, Yuchen Zhang, Yue Li, Yuning Mao, Zacharie Delpierre Coudert, Zheng Yan, Zhengxing Chen, Zoe Papakipos, Aaditya Singh, Aayushi Srivastava, Abha Jain, Adam Kelsey, Adam Shajnfeld, Adithya Gangidi, Adolfo Victoria, Ahuva Goldstand, Ajay Menon, Ajay Sharma, Alex Boesenberg, Alexei Baevski, Allie Feinstein, Amanda Kallet, Amit Sangani, Amos Teo, Anam Yunus, Andrei Lupu, Andres Alvarado, Andrew Caples, Andrew Gu, Andrew Ho, Andrew Poulton, Andrew Ryan, Ankit Ramchandani, Annie Dong, Annie Franco, Anuj Goyal, Aparajita Saraf, Arkabandhu Chowdhury, Ashley Gabriel, Ashwin Bharambe, Assaf Eisenman, Azadeh Yazdan, Beau James, Ben Maurer, Benjamin Leonhardi, Bernie Huang, Beth Loyd, Beto De Paola, Bhargavi Paranjape, Bing Liu, Bo Wu, Boyu Ni, Braden Hancock, Bram Wasti, Brandon Spence, Brani Stojkovic, Brian Gamido, Britt Montalvo, Carl Parker, Carly Burton, Catalina Mejia, Ce Liu, Changhan Wang, Changkyu Kim, Chao Zhou, Chester Hu, Ching-Hsiang Chu, Chris Cai, Chris Tindal, Christoph Feichtenhofer, Cynthia Gao, Damon Civin, Dana Beaty, Daniel Kreymer, Daniel Li, David Adkins, David Xu, Davide Testuggine, Delia David, Devi Parikh, Diana Liskovich, Didem Foss, Dingkang Wang, Duc Le, Dustin Holland, Edward Dowling, Eissa Jamil, Elaine Montgomery, Eleonora Presani, Emily Hahn, Emily Wood, Eric-Tuan Le, Erik Brinkman, Esteban Arcaute, Evan Dunbar, Evan Smothers, Fei Sun, Felix Kreuk, Feng Tian, Filippos Kokkinos, Firat Ozgenel, Francesco Caggioni, Frank Kanayet, Frank Seide, Gabriela Medina Florez, Gabriella Schwarz, Gada Badeer, Georgia Swee, Gil Halpern, Grant

Herman, Grigory Sizov, Guangyi, Zhang, Guna Lakshminarayanan, Hakan Inan, Hamid Shojanazeri, Han Zou, Hannah Wang, Hanwen Zha, Haroun Habeeb, Harrison Rudolph, Helen Suk, Henry Aspegren, Hunter Goldman, Hongyuan Zhan, Ibrahim Damlaj, Igor Molybog, Igor Tufanov, Ilias Leontiadis, Irina-Elena Veliche, Itai Gat, Jake Weissman, James Geboski, James Kohli, Janice Lam, Japhet Asher, Jean-Baptiste Gaya, Jeff Marcus, Jeff Tang, Jennifer Chan, Jenny Zhen, Jeremy Reizenstein, Jeremy Teboul, Jessica Zhong, Jian Jin, Jingyi Yang, Joe Cummings, Jon Carvill, Jon Shepard, Jonathan McPhie, Jonathan Torres, Josh Ginsburg, Junjie Wang, Kai Wu, Kam Hou U., Karan Saxena, Kartikay Khandelwal, Katayoun Zand, Kathy Matosich, Kaushik Veeraraghavan, Kelly Michelena, Keqian Li, Kiran Jagadeesh, Kun Huang, Kunal Chawla, Kyle Huang, Lailin Chen, Lakshya Garg, Lavender A., Leandro Silva, Lee Bell, Lei Zhang, Liangpeng Guo, Licheng Yu, Liron Moshkovich, Luca Wehrstedt, Madian Khabsa, Manav Avalani, Manish Bhatt, Martynas Mankus, Matan Hasson, Matthew Lennie, Matthias Reso, Maxim Groshev, Maxim Naumov, Maya Lathi, Meghan Keneally, Miao Liu, Michael L. Seltzer, Michal Valko, Michelle Restrepo, Mihir Patel, Mik Vyatskov, Mikayel Samvelyan, Mike Clark, Mike Macey, Mike Wang, Miquel Jubert Hermoso, Mo Metanat, Mohammad Rastegari, Munish Bansal, Nandhini Santhanam, Natascha Parks, Natasha White, Navyata Bawa, Nayan Singhal, Nick Egebo, Nicolas Usunier, Nikhil Mehta, Nikolay Pavlovich Laptev, Ning Dong, Norman Cheng, Oleg Chernoguz, Olivia Hart, Omkar Salpekar, Ozlem Kalinli, Parkin Kent, Parth Parekh, Paul Saab, Pavan Balaji, Pedro Rittner, Philip Bontrager, Pierre Roux, Piotr Dollar, Polina Zvyagina, Prashant Ratanchandani, Pritish Yuvraj, Qian Liang, Rachad Alao, Rachel Rodriguez, Rafi Ayub, Raghotham Murthy, Raghu Nayani, Rahul Mitra, Rangaprabhu Parthasarathy, Raymond Li, Rebekkah Hogan, Robin Battey, Rocky Wang, Russ Howes, Ruty Rinott, Sachin Mehta, Sachin Siby, Sai Jayesh Bondu, Samyak Datta, Sara Chugh, Sara Hunt, Sargun Dhillon, Sasha Sidorov, Satadru Pan, Saurabh Mahajan, Saurabh Verma, Seiji Yamamoto, Sharadh Ramaswamy, Shaun Lindsay, Shaun Lindsay, Sheng Feng, Shenghao Lin, Shengxin Cindy Zha, Shishir Patil, Shiva Shankar, Shuqiang Zhang, Shuqiang Zhang, Sinong Wang, Sneha Agarwal, Soji Sajuyigbe, Soumith Chintala, Stephanie Max, Stephen Chen, Steve Kehoe, Steve Satterfield, Sudarshan Govindaprasad, Sumit Gupta, Summer Deng, Sungmin Cho, Sunny Virk, Suraj Subramanian, Sy Choudhury, Sydney Goldman, Tal Remez, Tamar Glaser, Tamara Best, Thilo Koehler, Thomas Robinson, Tianhe Li, Tianjun Zhang, Tim Matthews, Timothy Chou, Tzook Shaked, Varun Vontimitta, Victoria Ajayi, Victoria Montanez, Vijai Mohan, Vinay Satish Kumar, Vishal Mangla, Vlad Ionescu, Vlad Poenaru, Vlad Tiberiu Mihailescu, Vladimir Ivanov, Wei Li, Wenchen Wang, Wenwen Jiang, Wes Bouaziz, Will Constable, Xiaocheng Tang, Xiaojian Wu, Xiaolan Wang, Xilun Wu, Xinbo Gao, Yaniv Kleinman, Yanjun Chen, Ye Hu, Ye Jia, Ye Qi, Yenda Li, Yilin Zhang, Ying Zhang, Yossi Adi, Youngjin Nam, Yu Wang, Yu Zhao, Yuchen Hao, Yundi Qian, Yunlu Li, Yuzi He, Zach Rait, Zachary DeVito, Zef Rosnbrick, Zhaoduo Wen, Zhenyu Yang, Zhiwei Zhao, and Zhiyu Ma. 2024. The Llama 3 herd of models. ArXiv:2407.21783 [cs].

Daniel Grießhaber, Ngoc Thang Vu, and Johannes Maucher. 2020. Low-resource text classification using domain-adversarial learning. *Computer Speech & Language*, 62:101056. https://doi.org/10.1016/j.csl.2019.101056

Himanshu Gupta, Saurabh Arjun Sawant, Swaroop Mishra, Mutsumi Nakamura, Arindam Mitra, Santosh Mashetty, and Chitta Baral. 2023. Instruction tuned models are quick learners. ArXiv:2306.05539 [cs].

Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith. 2020. Don't stop pretraining: Adapt language models to domains and tasks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8342–8360, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.740

Xu Han, Zhengyan Zhang, Ning Ding, Yuxian Gu, Xiao Liu, Yuqi Huo, Jiezhong Qiu, Yuan

Yao, Ao Zhang, Liang Zhang, Wentao Han, Minlie Huang, Qin Jin, Yanyan Lan, Yang Liu, Zhiyuan Liu, Zhiwu Lu, Xipeng Qiu, Ruihua Song, Jie Tang, Ji-Rong Wen, Jinhui Yuan, Wayne Xin Zhao, and Jun Zhu. 2021. Pre-trained models: Past, present and future. *AI Open*, 2:225–250. https://doi.org/10.1016/j.aiopen.2021.08.002

Zeyu Han, Chao Gao, Jinyang Liu, Jeff Zhang, and Sai Qian Zhang. 2024. Parameter-efficient fine-tuning for large models: A comprehensive survey. *Transactions on Machine Learning Research*.

Soufiane Hayou, Nikhil Ghosh, and Bin Yu. 2024. LoRA+: Efficient low rank adaptation of large models. In *Proceedings of the 41st International Conference on Machine Learning*, pages 17783–17806. PMLR.

Junxian He, Chunting Zhou, Xuezhe Ma, Taylor Berg-Kirkpatrick, and Graham Neubig. 2021. Towards a unified view of parameter-efficient transfer learning. In *International Conference on Learning Representations*.

Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. 2020. DeBERTa: Decoding-enhanced BERT with disentangled attention. In *International Conference on Learning Representations*.

Michael A. Hedderich, Lukas Lange, Heike Adel, Jannik Strötgen, and Dietrich Klakow. 2021. A survey on recent approaches for natural language processing in low-resource scenarios. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2545–2568, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.naacl-main.201

Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals, and Laurent Sifre. 2022. Training compute-optimal large language models. In *Advances in Neural Information Processing Systems*, volume 36, pages 30016–30030.

Jiwoo Hong, Noah Lee, and James Thorne. 2024. ORPO: Monolithic preference optimization without reference model. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 11170–11189, Miami, Florida, USA. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.626

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. 2019. Parameter-efficient transfer learning for NLP. In *Proceedings of the 36th International Conference on Machine Learning*, pages 2790–2799. PMLR.

Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 328–339, Melbourne, Australia. Association for Computational Linguistics. https://doi.org/10.18653/v1/P18-1031

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Xiyang Hu, Xinchi Chen, Peng Qi, Deguang Kong, Kunlun Liu, William Yang Wang, and Zhiheng Huang. 2023a. Language agnostic multilingual information retrieval with contrastive learning. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 9133–9146, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-acl.581

Zhiqiang Hu, Lei Wang, Yihuai Lan, Wanyu Xu, Ee-Peng Lim, Lidong Bing, Xing Xu, Soujanya Poria, and Roy Lee. 2023b. LLM-adapters: An

adapter family for parameter-efficient fine-tuning of large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5254–5276, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.319

Haoyang Huang, Tianyi Tang, Dongdong Zhang, Xin Zhao, Ting Song, Yan Xia, and Furu Wei. 2023a. Not all languages are created equal in LLMs: Improving multilingual capability by cross-lingual-thought prompting. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12365–12394, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.826

Yinqiu Huang, Min Gao, Jia Wang, Junwei Yin, Kai Shu, Qilin Fan, and Junhao Wen. 2023b. Meta-prompt based learning for low-resource false information detection. *Information Processing & Management*, 60(3):103279. https://doi.org/10.1016/j.ipm.2023.103279

Zhiqi Huang, Puxuan Yu, and James Allan. 2023c. Improving cross-lingual information retrieval on low-resource languages via optimal transport distillation. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, WSDM '23, pages 1048–1056, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3539597.3570468

Mathew Huerta-Enochian and Seung Yong Ko. 2024. Instruction fine-tuning: Does prompt loss matter? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 22771–22795, Miami, Florida, USA. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.1267

Chia-Chien Hung, Lukas Lange, and Jannik Strötgen. 2023. TADA: Efficient task-agnostic domain adaptation for Transformers. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 487–503, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-acl.31

Ayyoob Imani, Peiqin Lin, Amir Hossein Kargaran, Silvia Severini, Masoud Jalili Sabet, Nora Kassner, Chunlan Ma, Helmut Schmid, André Martins, François Yvon, and Hinrich Schütze. 2023. Glot500: Scaling multilingual corpora and language models to 500 languages. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1082–1117, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.61

Srinivasan Iyer, Xi Victoria Lin, Ramakanth Pasunuru, Todor Mihaylov, Daniel Simig, Ping Yu, Kurt Shuster, Tianlu Wang, Qing Liu, Punit Singh Koura, Xian Li, Brian O'Horo, Gabriel Pereyra, Jeff Wang, Christopher Dewan, Asli Celikyilmaz, Luke Zettlemoyer, and Ves Stoyanov. 2023. OPT-IML: Scaling language model instruction meta learning through the lens of generalization. ArXiv:2212.12017 [cs].

Kevin Maik Jablonka, Philippe Schwaller, Andres Ortega-Guerrero, and Berend Smit. 2024. Leveraging large language models for predictive chemistry. *Nature Machine Intelligence*, 6(2):161–169. Nature Publishing Group. https://doi.org/10.1038/s42256-023-00788-1

Neel Jain, Ping yeh Chiang, Yuxin Wen, John Kirchenbauer, Hong-Min Chu, Gowthami Somepalli, Brian R. Bartoldson, Bhavya Kailkhura, Avi Schwarzschild, Aniruddha Saha, Micah Goldblum, Jonas Geiping, and Tom Goldstein. 2024. NEFTune: Noisy embeddings improve instruction finetuning. In *The Twelfth International Conference on Learning Representations*.

Joel Jang, Seungone Kim, Seonghyeon Ye, Doyoung Kim, Lajanugen Logeswaran, Moontae Lee, Kyungjae Lee, and Minjoon Seo. 2023. Exploring the benefits of training expert language models over instruction tuning. In *Proceedings of the 40th International Conference on Machine Learning*, pages 14702–14729. PMLR.

Michael Jantscher, Felix Gunzer, Roman Kern, Eva Hassler, Sebastian Tschauner, and Gernot Reishofer. 2023. Information extraction from German radiological reports for general clinical

text and language understanding. *Scientific Reports*, 13(1):2353. Nature Publishing Group. https://doi.org/10.1038/s41598-023-29323-3, PubMed: 36759679

Ruili Jiang, Kehai Chen, Xuefeng Bai, Zhixuan He, Juntao Li, Muyun Yang, Tiejun Zhao, Liqiang Nie, and Min Zhang. 2024. A survey on human preference learning for large language models. ArXiv:2406.11191 [cs].

Hongpeng Jin, Wenqi Wei, Xuyu Wang, Wenbin Zhang, and Yanzhao Wu. 2023. Rethinking learning rate tuning in the era of large language models. In *2023 IEEE 5th International Conference on Cognitive Machine Intelligence (CogMI)*, pages 112–121. https://doi.org/10.1109/CogMI58952.2023.00025

Josip Jukić and Jan Snajder. 2023. Parameter-efficient language model tuning with active learning in low-resource settings. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5061–5074, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.307

Damjan Kalajdzievski. 2024. Scaling laws for forgetting when fine-tuning large language models. ArXiv:2401.05605v1 [cs].

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. 2020. Scaling laws for neural language models. ArXiv:2001.08361v1 [cs, stat].

Siddharth Karamcheti, Ranjay Krishna, Li Fei-Fei, and Christopher Manning. 2021. Mind your outliers! Investigating the negative impact of outliers on active learning for visual question answering. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7265–7281, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.564

Rabeeh Karimi Mahabadi, James Henderson, and Sebastian Ruder. 2021. Compacter: Efficient low-rank hypercomplex adapter layers. In *Advances in Neural Information Processing Systems*, volume 34, pages 1022–1035. Curran Associates, Inc.

Rabeeh Karimi Mahabadi, Luke Zettlemoyer, James Henderson, Lambert Mathias, Marzieh Saeidi, Veselin Stoyanov, and Majid Yazdani. 2022. Prompt-free and efficient few-shot learning with language models. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3638–3652, Dublin, Ireland. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.254

Payam Karisani. 2022. Multiple-source domain adaptation via coordinated domain encoders and paired classifiers. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(7):7087–7095. https://doi.org/10.1609/aaai.v36i7.20668

Md Kowsher, Md. Shohanur Islam Sobuj, Nusrat Jahan Prottasha, Mohammad Shamsul Arefin, and Yasuhiko Morimoto. 2023. Contrastive learning for universal zero-shot NLI with cross-lingual sentence embeddings. In *Proceedings of the 3rd Workshop on Multi-lingual Representation Learning (MRL)*, pages 239–252, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.mrl-1.18

Taku Kudo and John Richardson. 2018. SentencePiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 66–71, Brussels, Belgium. Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-2012, PubMed: 29382465

Guillaume Lample, Myle Ott, Alexis Conneau, Ludovic Denoyer, and Marc'Aurelio Ranzato. 2018. Phrase-based & neural unsupervised machine translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 5039–5049, Brussels, Belgium. Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-1549

Moritz Laurer, Wouter van Atteveldt, Andreu Casas, and Kasper Welbers. 2024. Less annotating, more classifying: Addressing the data scarcity issue of supervised machine learning with deep transfer learning and BERT-NLI. *Political Analysis*, 32(1):84–100. https://doi.org/10.1017/pan.2023.20

Cheolhyoung Lee, Kyunghyun Cho, and Wanmo Kang. 2019. Mixout: Effective regularization to finetune large-scale pretrained language models. In *International Conference on Learning Representations*.

Harrison Lee, Samrat Phatale, Hassan Mansoor, Thomas Mesnard, Johan Ferret, Kellie Ren Lu, Colton Bishop, Ethan Hall, Victor Carbune, Abhinav Rastogi, and Sushant Prakash. 2024. RLAIF vs. RLHF: Scaling reinforcement learning from human feedback with AI feedback. In *Proceedings of the 41st International Conference on Machine Learning*, pages 26874–26901. PMLR. https://doi.org/10.3384/VS.2001-5992.2024.11.1.1-8

Jens Lemmens and Walter Daelemans. 2023. Combining active learning and task adaptation with BERT for cost-effective annotation of social media datasets. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 237–250, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.wassa-1.22

Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. The power of scale for parameter-efficient prompt tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3045–3059, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.243

Jing Li, Billy Chiu, Shanshan Feng, and Hao Wang. 2022. Few-shot named entity recognition via meta-learning. *IEEE Transactions on Knowledge and Data Engineering*, 34(9):4245–4256. Conference Name: IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2020.3038670

Xiang Lisa Li and Percy Liang. 2021. Prefix-Tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4582–4597, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.353

Xuechen Li, Tianyi Zhang, Yann Dubois, Rohan Taori, Ishaan Gulrajani, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 2023. AlpacaEval: An automatic evaluator of instruction-following models.

Vladislav Lialin, Vijeta Deshpande, Xiaowei Yao, and Anna Rumshisky. 2024. Scaling down to scale up: A guide to parameter-efficient fine-tuning. ArXiv:2303.15647 [cs].

Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O'Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, and Xian Li. 2022. Few-shot learning with multilingual generative language models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9019–9052, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.616

Tongtao Ling, Lei Chen, Yutao Lai, and Hai-Lin Liu. 2023. Evolutionary verbalizer search for prompt-based few shot text classification. In *Knowledge Science, Engineering and Management*, pages 279–290, Cham. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-40292-0_23

Guanghui Liu. 2025. glgh/awesome-llm-human-preference-datasets. Original-date: 2023-05-03T16:48:10Z.

Hao Liu, Carmelo Sferrazza, and Pieter Abbeel. 2023a. Chain of hindsight aligns language models with feedback. In *International Conference on Learning Representations*.

Haokun Liu, Derek Tam, Mohammed Muqeeth, Jay Mohta, Tenghao Huang, Mohit Bansal, and

Colin A. Raffel. 2022. Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. *Advances in Neural Information Processing Systems*, 35:1950–1965.

Jia Liu, Yue Wang, Zhiqi Lin, Min Chen, Yixue Hao, and Long Hu. 2024a. Natural language fine-tuning. ArXiv:2412.20382 [cs].

Meizhen Liu, Xu Guo, He Jiakai, Jianye Chen, Fengyu Zhou, and Siu Hui. 2023b. InteMATs: Integrating granularity-specific multilingual adapters for cross-lingual transfer. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5035–5049, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.335

Meizhen Liu, Jiakai He, Xu Guo, Jianye Chen, Siu Cheung Hui, and Fengyu Zhou. 2023c. GranCATs: Cross-lingual enhancement through granularity-specific contrastive adapters. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, CIKM '23, pages 1461–1471, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3583780.3614896

Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023d. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35. https://doi.org/10.1145/3560815

Shih-Yang Liu, Chien-Yi Wang, Hongxu Yin, Pavlo Molchanov, Yu-Chiang Frank Wang, Kwang-Ting Cheng, and Min-Hung Chen. 2024b. DoRA: Weight-decomposed low-rank adaptation. In *Proceedings of the 41st International Conference on Machine Learning*, pages 32100–32121. PMLR.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. RoBERTa: A robustly optimized bert pretraining approach.

Shayne Longpre, Le Hou, Tu Vu, Albert Webson, Hyung Won Chung, Yi Tay, Denny Zhou, Quoc V. Le, Barret Zoph, Jason Wei, and Adam Roberts. 2023. The flan collection: Designing data and methods for effective instruction tuning. In *Proceedings of the 40th International Conference on Machine Learning*, pages 22631–22648. PMLR.

Keming Lu, Peter Potash, Xihui Lin, Yuwen Sun, Zihan Qian, Zheng Yuan, Tristan Naumann, Tianxi Cai, and Junwei Lu. 2023. Prompt discriminative language models for domain adaptation. In *Proceedings of the 5th Clinical Natural Language Processing Workshop*, pages 247–258, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.clinicalnlp-1.30

Wei Lu, Rachel K. Luu, and Markus J. Buehler. 2025. Fine-tuning large language models for domain adaptation: Exploration of training strategies, scaling, model merging and synergistic capabilities. *NPJ Computational Materials*, 11(1):1–43. Nature Publishing Group. https://doi.org/10.1038/s41524-025-01564-y

Aniruddha Mahapatra, Sharmila Reddy Nangi, Aparna Garimella, and Anandhavelu N. 2022. Entity extraction in low resource domains with selective pre-training of large language models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 942–951, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.61

Yuning Mao, Lambert Mathias, Rui Hou, Amjad Almahairi, Hao Ma, Jiawei Han, Scott Yih, and Madian Khabsa. 2022. UniPELT: A unified framework for parameter-efficient language model tuning. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6253–6264, Dublin, Ireland. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.433

Yu Meng, Mengzhou Xia, and Danqi Chen. 2024. SimPO: Simple preference optimization with a reference-free reward. *Advances in Neural Information Processing Systems*,

37:124198–124235. https://doi.org/10.52202/079017-3946

Meta AI. 2025. The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation.

Tomas Mikolov, Quoc V. Le, and Ilya Sutskever. 2013. Exploiting similarities among languages for machine translation. ArXiv:1309.4168v1 [cs].

Sewon Min, Mike Lewis, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2022. MetaICL: Learning to learn in context. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2791–2809, Seattle, United States. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.201

Benjamin Minixhofer, Fabian Paischer, and Navid Rekabsaz. 2022. WECHSEL: Effective initialization of subword embeddings for cross-lingual transfer of monolingual language models. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3992–4006, Seattle, United States. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.naacl-main.293

Vincenzo Moscato, Giuseppe Napolano, Marco Postiglione, and Giancarlo Sperlì. 2023. Multi-task learning for few-shot biomedical relation extraction. *Artificial Intelligence Reviews*, 56(11):13743–13763. https://doi.org/10.1007/s10462-023-10484-6

Niklas Muennighoff, Alexander Rush, Boaz Barak, Teven Le Scao, Nouamane Tazi, Aleksandra Piktus, Sampo Pyysalo, Thomas Wolf, and Colin A. Raffel. 2023a. Scaling data-constrained language models. *Advances in Neural Information Processing Systems*, 36:50358–50376.

Niklas Muennighoff, Thomas Wang, Lintang Sutawika, Adam Roberts, Stella Biderman, Teven Le Scao, M. Saiful Bari, Sheng Shen, Zheng Xin Yong, Hailey Schoelkopf, Xiangru Tang, Dragomir Radev, Alham Fikri Aji, Khalid Almubarak, Samuel Albanie, Zaid Alyafeai, Albert Webson, Edward Raff, and Colin Raffel. 2023b. Crosslingual generalization through multitask finetuning. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15991–16111, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.891

Mohammed Muqeeth, Haokun Liu, Yufan Liu, and Colin Raffel. 2024. Learning to route among specialized experts for zero-shot generalization. In *International Conference on Machine Learning*.

Saliha Muradoglu and Mans Hulden. 2022. Eeny, meeny, miny, moe. How to choose data for morphological inflection. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 7294–7303, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.492

Arijit Nag, Bidisha Samanta, Animesh Mukherjee, Niloy Ganguly, and Soumen Chakrabarti. 2023. Entropy-guided vocabulary augmentation of multilingual language models for low-resource tasks. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 8619–8629, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-acl.548

Nghia Ngo Trung, Duy Phung, and Thien Huu Nguyen. 2021. Unsupervised domain adaptation for event detection using domain-specific adapters. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4015–4025, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.findings-acl.351

Mariana Noguti, Eduardo Vellasques, and Luiz S. Oliveira. 2023. A small claims court for the NLP: Judging legal text classification strategies with small datasets. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1840–1845. https://doi.org/10.1109/SMC53992.2023.10394189

OpenAI, Sandhini Agarwal, Lama Ahmad, Jason Ai, Sam Altman, Andy Applebaum, Edwin Arbus, Rahul K. Arora, Yu Bai, Bowen Baker, Haiming Bao, Boaz Barak, Ally Bennett, Tyler Bertao, Nivedita Brett, Eugene Brevdo, Greg Brockman, Sebastien Bubeck, Che Chang, Kai Chen, Mark Chen, Enoch Cheung, Aidan Clark, Dan Cook, Marat Dukhan, Casey Dvorak, Kevin Fives, Vlad Fomenko, Timur Garipov, Kristian Georgiev, Mia Glaese, Tarun Gogineni, Adam Goucher, Lukas Gross, Katia Gil Guzman, John Hallman, Jackie Hehir, Johannes Heidecke, Alec Helyar, Haitang Hu, Romain Huet, Jacob Huh, Saachi Jain, Zach Johnson, Chris Koch, Irina Kofman, Dominik Kundel, Jason Kwon, Volodymyr Kyrylov, Elaine Ya Le, Guillaume Leclerc, James Park Lennon, Scott Lessans, Mario Lezcano-Casado, Yuanzhi Li, Zhuohan Li, Ji Lin, Jordan Liss, Lily, Liu, Jiancheng Liu, Kevin Lu, Chris Lu, Zoran Martinovic, Lindsay McCallum, Josh McGrath, Scott McKinney, Aidan McLaughlin, Song Mei, Steve Mostovoy, Tong Mu, Gideon Myles, Alexander Neitz, Alex Nichol, Jakub Pachocki, Alex Paino, Dana Palmie, Ashley Pantuliano, Giambattista Parascandolo, Jongsoo Park, Leher Pathak, Carolina Paz, Ludovic Peran, Dmitry Pimenov, Michelle Pokrass, Elizabeth Proehl, Huida Qiu, Gaby Raila, Filippo Raso, Hongyu Ren, Kimmy Richardson, David Robinson, Bob Rotsted, Hadi Salman, Suvansh Sanjeev, Max Schwarzer, D. Sculley, Harshit Sikchi, Kendal Simon, Karan Singhal, Yang Song, Dane Stuckey, Zhiqing Sun, Philippe Tillet, Sam Toizer, Foivos Tsimpourlas, Nikhil Vyas, Eric Wallace, Xin Wang, Miles Wang, Olivia Watkins, Kevin Weil, Amy Wendling, Kevin Whinnery, Cedric Whitney, Hannah Wong, Lin Yang, Yu Yang, Michihiro Yasunaga, Kristen Ying, Wojciech Zaremba, Wenting Zhan, Cyril Zhang, Brian Zhang, Eddie Zhang, and Shengjia Zhao. 2025. gpt-oss-120b & gpt-oss-20b model card. ArXiv:2508.10925 [cs].

Oleksiy Ostapenko, Zhan Su, Edoardo Ponti, Laurent Charlin, Nicolas Le Roux, Lucas Caccia, and Alessandro Sordoni. 2024. Towards modular LLMs by building and reusing a library of LoRAs. In *Proceedings of the 41st International Conference on Machine Learning*, pages 38885–38904. PMLR.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

Aldo Pareja, Nikhil Shivakumar Nayak, Hao Wang, Krishnateja Killamsetty, Shivchander Sudalairaj, Wenlong Zhao, Seungwook Han, Abhishek Bhandwaldar, Guangxuan Xu, Kai Xu, Ligong Han, Luke Inglis, and Akash Srivastava. 2024. Unveiling the secret recipe: A guide for supervised fine-tuning small LLMs. ArXiv:2412.13337 [cs].

Ryan Park, Rafael Rafailov, Stefano Ermon, and Chelsea Finn. 2024. Disentangling length from quality in direct preference optimization. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 4998–5017, Bangkok, Thailand. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-acl.297, PubMed: 38866830

Amalie Pauli, Leon Derczynski, and Ira Assent. 2023. Anchoring fine-tuning of sentence transformer with semantic label information for efficient truly few-shot classification. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 11254–11264, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.692

Letian Peng, Zihan Wang, and Jingbo Shang. 2023. Less than one-shot: Named entity recognition via extremely weak supervision. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 13603–13616, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.908, PubMed: 37279247

Jonas Pfeiffer, Aishwarya Kamath, Andreas Rücklé, Kyunghyun Cho, and Iryna Gurevych. 2021. AdapterFusion: Non-destructive task

composition for transfer learning. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 487–503, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.eacl-main.39

Jonas Pfeiffer, Sebastian Ruder, Ivan Vulić, and Edoardo Ponti. 2023. Modular deep learning. *Transactions on Machine Learning Research*.

Jonas Pfeiffer, Ivan Vulić, Iryna Gurevych, and Sebastian Ruder. 2020. MAD-X: An adapter-based framework for multi-task cross-lingual transfer. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7654–7673, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.617

Jason Phang, Thibault Févry, and Samuel R. Bowman. 2019. Sentence encoders on STILTs: Supplementary training on intermediate labeled-data tasks. ArXiv:1811.01088v2 [cs].

Edoardo M. Ponti, Alessandro Sordoni, Yoshua Bengio, and Siva Reddy. 2022. Combining modular skills in multitask learning. ArXiv: 2202.13914 [cs].

Yada Pruksachatkun, Jason Phang, Haokun Liu, Phu Mon Htut, Xiaoyi Zhang, Richard Yuanzhe Pang, Clara Vania, Katharina Kann, and Samuel R. Bowman. 2020. Intermediate-task transfer learning with pretrained language models: When and why does it work? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5231–5247, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.467

Kunxun Qi, Hai Wan, Jianfeng Du, and Haolan Chen. 2022. Enhancing cross-lingual natural language inference by prompt-learning from cross-lingual templates. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1910–1923, Dublin, Ireland. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.134

Libo Qin, Qiguang Chen, Fuxuan Wei, Shijue Huang, and Wanxiang Che. 2023. Cross-lingual prompting: Improving zero-shot chain-of-thought reasoning across languages. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2695–2709, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.163

Muhammad Qorib, Geonsik Moon, and Hwee Tou Ng. 2024. Are decoder-only language models better than encoder-only language models in understanding word meaning? In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 16339–16347, Bangkok, Thailand. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-acl.967

Rafael Rafailov, Yaswanth Chittepu, Ryan Park, Harshit Sikchi, Joey Hejna, W. B. Knox, Chelsea Finn, and Scott Niekum. 2024. Scaling laws for reward model overoptimization in direct alignment algorithms. *Advances in Neural Information Processing Systems*, 37:126207–126242. https://doi.org/10.52202/079017-4009

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36:53728–53741.

Rajkumar Ramamurthy, Prithviraj Ammanabrolu, Kianté Brantley, Jack Hessel, Rafet Sifa, Christian Bauckhage, Hannaneh Hajishirzi, and Yejin Choi. 2022. Is reinforcement learning (not) for natural language processing: Benchmarks, baselines, and building blocks for natural language policy optimization. In *International Conference on Learning Representations*.

Vinay Venkatesh Ramasesh, Aitor Lewkowycz, and Ethan Dyer. 2021. Effect of scale on catastrophic forgetting in neural networks. In *International Conference on Learning Representations*.

Andreas Rücklé, Gregor Geigle, Max Glockner, Tilman Beck, Jonas Pfeiffer, Nils Reimers, and Iryna Gurevych. 2021. AdapterDrop: On the efficiency of adapters in transformers. In

Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pages 7930–7946, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.626

Shuo Ren, Zhirui Zhang, Shujie Liu, Ming Zhou, and Shuai Ma. 2019. Unsupervised neural machine translation with SMT as posterior regularization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1):241–248. https://doi.org/10.1609/aaai.v33i01.3301241

Sebastian Ruder, Jonathan Clark, Alexander Gutkin, Mihir Kale, Min Ma, Massimo Nicosia, Shruti Rijhwani, Parker Riley, Jean-Michel Sarr, Xinyi Wang, John Wieting, Nitish Gupta, Anna Katanova, Christo Kirov, Dana Dickinson, Brian Roark, Bidisha Samanta, Connie Tao, David Adelani, Vera Axelrod, Isaac Caswell, Colin Cherry, Dan Garrette, Reeve Ingle, Melvin Johnson, Dmitry Panteleev, and Partha Talukdar. 2023. XTREME-UP: A user-centric scarce-data benchmark for under-represented languages. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1856–1884, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.125

Hossain Shaikh Saadi, Viktor Hangya, Tobias Eder, and Alexander Fraser. 2022. Comparative analysis of cross-lingual contextualized word embeddings. In *Proceedings of the 2nd Workshop on Multi-lingual Representation Learning (MRL)*, pages 64–75, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.mrl-1.6

Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M. Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Tali Bers, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M. Rush. 2022. Multitask prompted training enables zero-shot task generalization. In *International Conference on Learning Representations*.

Timo Schick and Hinrich Schütze. 2021a. Exploiting cloze-questions for few-shot text classification and natural language inference. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 255–269, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.eacl-main.20

Timo Schick and Hinrich Schütze. 2021b. It's not just size that matters: Small language models are also few-shot learners. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2339–2352, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.naacl-main.185

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. ArXiv:1707.06347 [cs].

Archit Sharma, Sedrick Keh, Eric Mitchell, Chelsea Finn, Kushal Arora, and Thomas Kollar. 2024. A critical evaluation of AI feedback for aligning large language models. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. 2017. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*.

Judy Hanwen Shen, Archit Sharma, and Jun Qin. 2024. Towards data-centric RLHF: Simple metrics for preference dataset comparison. ArXiv:2409.09603 [cs].

Zhengyan Shi, Adam X. Yang, Bin Wu, Laurence Aitchison, Emine Yilmaz, and Aldo Lipani.

2024. Instruction tuning with loss over instructions. *Advances in Neural Information Processing Systems*, 37:69176–69205. https://doi.org/10.52202/079017-2210

Ken Shoemake. 1985. Animating rotation with quaternion curves. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '85, pages 245–254, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/325165.325242

Jake Snell, Kevin Swersky, and Richard Zemel. 2017. Prototypical networks for few-shot learning. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. 2020. FixMatch: Simplifying semi-supervised learning with consistency and confidence. *Advances in Neural Information Processing Systems*, 33:596–608.

Chengyu Song, Fei Cai, Mengru Wang, Jianming Zheng, and Taihua Shao. 2023. TaxonPrompt: Taxonomy-aware curriculum prompt learning for few-shot event classification. *Knowledge-Based Systems*, 264:110290. https://doi.org/10.1016/j.knosys.2023.110290

Nikolaos Stylianou, Despoina Chatzakou, Theodora Tsikrika, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2023. Domain-aligned data augmentation for low-resource and imbalanced text classification. In *Advances in Information Retrieval*, pages 172–187, Cham. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-28238-6_12

Yi-Lin Sung, Jaemin Cho, and Mohit Bansal. 2022. LST: Ladder side-tuning for parameter and memory efficient transfer learning. *Advances in Neural Information Processing Systems*, 35:12991–13005.

Yi-Lin Sung, Varun Nair, and Colin A. Raffel. 2021. Training neural networks with fixed sparse masks. In *Advances in Neural Information Processing Systems*, volume 34, pages 24193–24205. Curran Associates, Inc.

Wen Tai, H. T. Kung, Xin Dong, Marcus Comiter, and Chang-Fu Kuo. 2020. exBERT: Extending pre-trained models with domain-specific vocabulary under constrained training resources. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1433–1439, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.findings-emnlp.129

Mingxu Tao, Chen Zhang, Quzhe Huang, Tianyao Ma, Songfang Huang, Dongyan Zhao, and Yansong Feng. 2024. Unlocking the potential of model merging for low-resource languages. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 8705–8720, Miami, Florida, USA. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-emnlp.508

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models. ArXiv:2307.09288 [cs].

Marcos Treviso, Ji-Ung Lee, Tianchu Ji, Betty van Aken, Qingqing Cao, Manuel R. Ciosici, Michael Hassid, Kenneth Heafield, Sara

Hooker, Colin Raffel, Pedro H. Martins, André F. T. Martins, Jessica Zosa Forde, Peter Milder, Edwin Simpson, Noam Slonim, Jesse Dodge, Emma Strubell, Niranjan Balasubramanian, Leon Derczynski, Iryna Gurevych, and Roy Schwartz. 2023. Efficient methods for natural language processing: A survey. *Transactions of the Association for Computational Linguistics*, 11:826–860. https://doi.org/10.1162/tacl_a_00577

Hazal Türkmen, Oguz Dikenelli, Cenk Eraslan, Mehmet Cem Çalli, and Süha Süreyya Özbek. 2023. BioBERTurk: Exploring turkish biomedical language model development strategies in low-resource setting. *Journal of Healthcare Informatics Research*, 7(4):433–446. https://doi.org/10.1007/s41666-023-00140-7, PubMed: 37927378

Lifu Tu, Jin Qu, Semih Yavuz, Shafiq Joty, Wenhao Liu, Caiming Xiong, and Yingbo Zhou. 2024. Efficiently aligned cross-lingual transfer learning for conversational tasks using prompt-tuning. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 1278–1294, St. Julian's, Malta. Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-eacl.85

Lewis Tunstall, Edward Emanuel Beeching, Nathan Lambert, Nazneen Rajani, Kashif Rasul, Younes Belkada, Shengyi Huang, Leandro Von Werra, Clémentine Fourrier, Nathan Habib, Nathan Sarrazin, Omar Sanseviero, Alexander M. Rush, and Thomas Wolf. 2024. Zephyr: Direct distillation of LM alignment. In *First Conference on Language Modeling*.

Faizad Ullah, Ubaid Azam, Ali Faheem, Faisal Kamiran, and Asim Karim. 2023. Comparing prompt-based and standard fine-tuning for urdu text classification. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 6747–6754, Singapore. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.449

Dave Van Veen, Cara Van Uden, Maayane Attias, Anuj Pareek, Christian Bluethgen, Malgorzata Polacin, Wah Chiu, Jean-Benoit Delbrouck, Juan Zambrano Chaves, Curtis Langlotz, Akshay Chaudhari, and John Pauly. 2023. RadAdapt: Radiology report summarization via lightweight domain adaptation of large language models. In *The 22nd Workshop on Biomedical Natural Language Processing and BioNLP Shared Tasks*, pages 449–460, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.bionlp-1.42

Leandro von Werra, Younes Belkada, Lewis Tunstall, Edward Beeching, Tristan Thrush, Nathan Lambert, Shengyi Huang, Kashif Rasul, and Quentin Gallouédec. 2020. TRL: Transformer reinforcement learning. https://github.com/huggingface/trl.

Tu Vu, Tong Wang, Tsendsuren Munkhdalai, Alessandro Sordoni, Adam Trischler, Andrew Mattarella-Micke, Subhransu Maji, and Mohit Iyyer. 2020. Exploring and predicting transferability across NLP tasks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7882–7926, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.635

Zhongwei Wan, Xin Wang, Che Liu, Samiul Alam, Yu Zheng, Jiachen Liu, Zhongnan Qu, Shen Yan, Yi Zhu, Quanlu Zhang, Mosharaf Chowdhury, and Mi Zhang. 2024. Efficient large language models: A survey. *Transactions on Machine Learning Research*.

Jiaan Wang, Fandong Meng, Duo Zheng, Yunlong Liang, Zhixu Li, Jianfeng Qu, and Jie Zhou. 2023a. Towards unifying multi-lingual and cross-lingual summarization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15127–15143. Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.843

Minzheng Wang, Jia Cao, Qingchao Kong, and Yin Luo. 2023b. Boosting domain-specific question answering through weakly supervised self-training. In *2023 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 1–6. https://doi.org/10.1109/ISI58743.2023.10297258

Xiao Wang, Weikang Zhou, Can Zu, Han Xia, Tianze Chen, Yuansen Zhang, Rui Zheng, Junjie Ye, Qi Zhang, Tao Gui, Jihua Kang, Jingsheng Yang, Siyuan Li, and Chunsai Du. 2023c. InstructUIE: Multi-task instruction tuning for unified information extraction. ArXiv:2304.08085 [cs].

Yaqing Wang, Sahaj Agarwal, Subhabrata Mukherjee, Xiaodong Liu, Jing Gao, Ahmed Hassan Awadallah, and Jianfeng Gao. 2022. AdaMix: Mixture-of-adaptations for parameter-efficient model tuning. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5744–5760, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.emnlp-main.388

Yizhong Wang, Hamish Ivison, Pradeep Dasigi, Jack Hessel, Tushar Khot, Khyathi Chandu, David Wadden, Kelsey MacMillan, Noah A. Smith, Iz Beltagy, and Hannaneh Hajishirzi. 2023d. How far can camels go? Exploring the state of instruction tuning on open resources. *Advances in Neural Information Processing Systems*, 36:74764–74786.

Jason Wei, Maarten Bosma, Vincent Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. 2021. Finetuned language models are zero-shot learners. In *International Conference on Learning Representations*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.

Wen Wen, Yongbin Liu, Chunping Ouyang, Qiang Lin, and Tonglee Chung. 2021. Enhanced prototypical network for few-shot relation extraction. *Information Processing & Management*, 58(4):102596. https://doi.org/10.1016/j.ipm.2021.102596

Georg Wiese, Dirk Weissenborn, and Mariana Neves. 2017. Neural domain adaptation for biomedical question answering. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages 281–289, Vancouver, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/K17-1029

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-demos.6

Mitchell Wortsman, Gabriel Ilharco, Samir Ya Gadre, Rebecca Roelofs, Raphael Gontijo-Lopes, Ari S. Morcos, Hongseok Namkoong, Ali Farhadi, Yair Carmon, Simon Kornblith, and Ludwig Schmidt. 2022. Model soups: Averaging weights of multiple fine-tuned models improves accuracy without increasing inference time. In *Proceedings of the 39th International Conference on Machine Learning*, pages 23965–23998. PMLR.

Hui Wu, Yuting He, Yidong Chen, Yu Bai, and Xiaodong Shi. 2024. Improving few-shot relation extraction through semantics-guided learning. *Neural Networks*, 169:453–461. https://doi.org/10.1016/j.neunet.2023.10.053, PubMed: 37939534

Ting-Wei Wu, Changsheng Zhao, Ernie Chang, Yangyang Shi, Pierce Chuang, Vikas Chandra, and Biing Juang. 2023. Towards zero-shot multilingual transfer for code-switched responses. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7551–7563, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.417

Mengzhou Xia, Sadhika Malladi, Suchin Gururangan, Sanjeev Arora, and Danqi Chen. 2024. LESS: Selecting influential data for

targeted instruction tuning. In *Proceedings of the 41st International Conference on Machine Learning*, pages 54104–54132. PMLR.

Qizhe Xie, Zihang Dai, Eduard Hovy, Thang Luong, and Quoc Le. 2020. Unsupervised data augmentation for consistency training. In *Advances in Neural Information Processing Systems*, volume 33, pages 6256–6268. Curran Associates, Inc.

Sang Michael Xie, Shibani Santurkar, Tengyu Ma, and Percy S. Liang. 2023. Data selection for language models via importance resampling. *Advances in Neural Information Processing Systems*, 36:34201–34227.

Haoran Xu, Amr Sharaf, Yunmo Chen, Weiting Tan, Lingfeng Shen, Benjamin Van Durme, Kenton Murray, and Young Jin Kim. 2024. Contrastive preference optimization: Pushing the boundaries of LLM performance in machine translation. In *Proceedings of the 41st International Conference on Machine Learning*, pages 55204–55224. PMLR.

Mengwei Xu, Dongqi Cai, Wangsong Yin, Shangguang Wang, Xin Jin, and Xuanzhe Liu. 2025. Resource-efficient algorithms and systems of foundation models: A survey. *ACM Computing Surveys*, 57(5):110:1–110:39. https://doi.org/10.1145/3706418

Peng Xu, Dhruv Kumar, Wei Yang, Wenjie Zi, Keyi Tang, Chenyang Huang, Jackie Chi Kit Cheung, Simon J. D. Prince, and Yanshuai Cao. 2021a. Optimizing deeper transformers on small datasets. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 2089–2102, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.163

Runxin Xu, Fuli Luo, Zhiyuan Zhang, Chuanqi Tan, Baobao Chang, Songfang Huang, and Fei Huang. 2021b. Raise a child in large language model: Towards effective and generalizable fine-tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9514–9528, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.749

Yuanmeng Yan, Rumei Li, Sirui Wang, Fuzheng Zhang, Wei Wu, and Weiran Xu. 2021. ConSERT: A contrastive framework for self-supervised sentence representation transfer. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5065–5075, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.393

Charles Yu, Sullam Jeoung, Anish Kasi, Pengfei Yu, and Heng Ji. 2023. Unlearning bias in language models by partitioning gradients. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 6032–6048. Toronto, Canada. Association for Computational Linguistics.

Michelle Yuan, Hsuan-Tien Lin, and Jordan Boyd-Graber. 2020. Cold-start active learning through self-supervised language modeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7935–7948, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.637

Zheng Yuan, Hongyi Yuan, Chuanqi Tan, Wei Wang, Songfang Huang, and Fei Huang. 2023. RRHF: Rank responses to align language models with human feedback without tears. ArXiv:2304.05302 [cs].

Biao Zhang, Zhongtao Liu, Colin Cherry, and Orhan Firat. 2023a. When scaling meets LLM finetuning: The effect of data, model and finetuning method. In *The Twelfth International Conference on Learning Representations*.

Haojie Zhang, Ge Li, Jia Li, Zhongjin Zhang, Yuqi Zhu, and Zhi Jin. 2022a. Fine-tuning pre-trained language models effectively by optimizing subnetworks adaptively. *Advances in Neural Information Processing Systems*, 35:21442–21454.

Qingru Zhang, Minshuo Chen, Alexander Bukharin, Pengcheng He, Yu Cheng, Weizhu Chen, and Tuo Zhao. 2022b. Adaptive budget allocation for parameter-efficient fine-tuning. In *The Eleventh International Conference on Learning Representations*.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–595. https://doi.org/10.1109/CVPR.2018.00068

Shengyu Zhang, Linfeng Dong, Xiaoya Li, Sen Zhang, Xiaofei Sun, Shuhe Wang, Jiwei Li, Runyi Hu, Tianwei Zhang, Fei Wu, and Guoyin Wang. 2024. Instruction tuning for large language models: A survey. ArXiv:2308.10792 [cs].

Tianyi Zhang, Felix Wu, Arzoo Katiyar, Kilian Q. Weinberger, and Yoav Artzi. 2021. Revisiting few-sample BERT fine-tuning. In *International Conference on Learning Representations*.

Yue Zhang, Leyang Cui, Deng Cai, Xinting Huang, Tao Fang, and Wei Bi. 2023b. Multi-task instruction tuning of LLaMa for specific scenarios: A preliminary study on writing assistance. ArXiv:2305.13225 [cs].

Haodong Zhao, Ruifang He, Mengnan Xiao, and Jing Xu. 2023a. Infusing hierarchical guidance into prompt tuning: A parameter-efficient framework for multi-level implicit discourse relation recognition. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6477–6492, Toronto, Canada. Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.acl-long.357

Lingyun Zhao, Lin Li, Xinhao Zheng, and Jianwei Zhang. 2021. A BERT based sentiment analysis and key entity detection approach for online financial texts. In *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 1233–1238. https://doi.org/10.1109/CSCWD49262.2021.9437616

Lulu Zhao, Fujia Zheng, Weihao Zeng, Keqing He, Ruotong Geng, Huixing Jiang, Wei Wu, and Weiran Xu. 2022. ADPL: Adversarial prompt-based domain adaptation for dialogue summarization with knowledge disentanglement. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '22, pages 245–255, New York, NY, USA. Association for Computing Machinery. https://doi.org/10.1145/3477495.3531933

Yao Zhao, Rishabh Joshi, Tianqi Liu, Misha Khalman, Mohammad Saleh, and Peter J. Liu. 2023b. SLiC-HF: Sequence likelihood calibration with human feedback. ArXiv:2305.10425 [cs].

Chunting Zhou, Pengfei Liu, Puxin Xu, Srinivasan Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, Lili Yu, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, and Omer Levy. 2023. LIMA: Less is more for alignment. *Advances in Neural Information Processing Systems*, 36:55006–55021.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2223–2232.

Daniel M. Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B. Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. 2020. Fine-tuning language models from human preferences. ArXiv:1909.08593 [cs].

A Task-specific Training Methods

This section serves as justification and further contextualization for the task-specific suggestions in Table 4.

A.1 Text Classification

In ultra-low-resource settings (2–8 examples per class), contrastive learning (CL) methods like few-shot sentence Transformers trained with triplet loss and anchoring on class descriptions are effective (Pauli et al., 2023). Reformulating

Task	Continued Pre-training	Labeled Data*	Fine-tuning	Complementary Options
Text Classification	75K+ tokens for PEFT params, LR warm-up + ER	< 250	PET + Adapters	Label embeddings; Similar task data
		< 1K	NLI format with Prefix-tuning / UniPELT	Fine-tune first on high-resource NLI data
		< 10K	LoRA; LLRD	Active or Semi-supervised learning
		10K+	Full fine-tuning	Regularization (small batches, gradient clipping)
Sequence Labeling	25K+ tokens for PEFT params, LR warm-up + ER + Mixout	< 250	PET; QA format with CL	Intermediate training, CL between query/entity
		< 5K	Adapters, Parameter masking	Similar task datasets for knowledge distillation, Active learning
		5K+	Full fine-tuning	LLRD; Bounded gradients
Text Comparison	75K+ tokens for PEFT params, consistency CL, LR warm-up	< 250	PET + Adapters / Prompt-tuning	CL with label embeddings, Data augmentation
		< 1K	Adapters / UniPELT	Multilingual data with CL
		< 10K	PEFT (LoRA)	CL on NLI data
		10K+	Full fine-tuning	Bounded gradients

Table 4: Suggested approaches for NLP task groups with limited data. All information in this table was compiled from reviewed papers (Appendix A). Experience replay (ER) stands for mixing in some general-domain data as regularization. *Labeled data for Sequence Labeling is provided in sentences; for other tasks, in examples.

classification as NLI improves few-shot performance by enabling transfer from general-purpose NLI datasets (Laurer et al., 2024), and is especially effective with <1k examples. For similar data sizes, pattern-exploiting training (PET) consistently outperforms full FT across models like RoBERTa, XLM-R, and DistilBERT on NLI tasks, topic/sentiment classification (Schick and Schütze, 2021a; Ullah et al., 2023; Ling et al., 2023), and can be improved by using label embeddings and adapters (Karimi Mahabadi et al., 2022).

For $<$1k examples, prefix tuning and UniPELT outperform full FT with BERT, especially when combined with CPT on unlabeled data (75k tokens) (Jukić and Snajder, 2023). Soft prompt tuning is also superior in few-shot settings (2–64 shots), particularly with domain-specific models like ClinicalBERT and SciBERT (Goswami et al., 2023). Domain adaptation (CPT) with token millions, including vocabulary injection and last-layer tuning, further boosts performance on document classification (Gnehm et al., 2022). Prefix tuning works well up to 1k samples, while LoRA improves between 1k–10k, and full FT becomes competitive beyond that (Mao et al., 2022). Regularization methods like layerwise learning rate decay (LLRD) and mixout help stabilize training with BERT on GLUE tasks under 10k examples (Zhang et al., 2021). Active learning becomes impactful above 1k samples and can reduce labeling needs by up to 70% (Lemmens and Daelemans, 2023; Ein-Dor et al., 2020; Yuan et al., 2020).

For higher-resource setups ($\sim$10k examples), full FT performs best, especially when paired with gradient clipping and small batch sizes for stability in reading comprehension (Xu et al., 2021a). Finally, for tasks like NLI, even a subset of 16k examples (1.9M tokens) selected through clustering and coreset sampling is sufficient to train competitive decoder models like Galactica 1.3B (Chen et al., 2023a).

A.2 Sequence Labeling

In extreme low-resource setups (1–10 examples), pseudo-labeling enables 1-shot NER by using predicted spans as supervision (Peng et al., 2023). Reformulating NER as QA and applying CL between the query and target entity (positive pair) can further enhance few-shot learning (5–50 shots) in biomedical NER. Transferring knowledge from high-resource QA tasks boosts performance (Chen et al., 2023c). Similarly, for biomedical relation

extraction (RE), intermediate training with multiple related datasets and knowledge distillation from an ensemble outperforms PET, especially when task similarity (e.g., vocabulary overlap) is high (Moscato et al., 2023). In multilingual RE, mT5 (220M) underperforms XLM-R base (125M) despite its size, especially in languages with less pretraining data. For low-resource scenarios (<32 examples per class), prompts and label names should be kept in English for optimal transfer (Chen et al., 2022b). At 100–1k sentences, CPT domain adaptation is particularly effective. Using as few as 36k tokens improves BERT-based NER, and selecting the most relevant data subset yields better performance than using full corpora (Mahapatra et al., 2022). In German clinical RE/NER, FT with just 512–1400 sentences is feasible when combined with active learning (perplexity-based sampling) and domain-specific models like German-MedBERT or R-BERT (Jantscher et al., 2023). In the 1k–5k sentence range, PEFT strategies remain effective. CPT domain-adaptation offers limited gains, as fine-tuning the base model sometimes performs equally well (Crema et al., 2023). Additional techniques like LLRD, mixout, and learning rate warm-up improve generalization for NER, RE, and QA across tasks under 8k sentences (Buonocore et al., 2023). Adapter-based extensions with limited FT (∼5k sentences) also prove competitive when pretrained on large domain-specific corpora (Tai et al., 2020).

At moderate scale (5k–10k), BERT-based models trained from scratch on large biomedical corpora (12B tokens) perform comparably to those adapted via CPT on 1B tokens, offering flexibility in resource usage (Bai et al., 2021).

A.3 Text Comparison

CL proves to be a powerful pretraining strategy for improving semantic representations in tasks like NLI or paraphrase detection. Models like DeBERTa large (435M) benefit significantly from CL pretraining when used for zero- or few-shot textual entailment (Kowsher et al., 2023). Similarly, SimCSE and ConSERT show that both unsupervised CL and supervised NLI-based learning greatly enhance sentence embeddings without labeled task data (Gao et al., 2021; Yan et al., 2021). For ultra-low supervision (16–64 examples), PEFT techniques consistently outperform full FT. RoBERTa large (355M), using a cloze-style format with adapters and label embedding learning, surpasses FT, BitFit, PET, and prompt tuning with just 16 samples per label (Karimi Mahabadi et al., 2022). PET also performs well for multi-level implicit discourse relation recognition using RoBERTa (Zhao et al., 2023a). In multilingual and cross-lingual comparison tasks, PET combined with CL consistency loss shows strong results. On tasks like XNLI and paraphrase identification, models such as XLM-R base (279M), XLM-R large (561M), and InfoXLM large (561M) benefit from this hybrid approach, especially in 16–256-shot learning settings (Qi et al., 2022). A CL framework using 500 parallel sentence pairs, supplemented with up to 10k non-parallel instances, establishes strong information retrieval performance across languages (Hu et al., 2023a). For 1k–10k instances, lightweight tuning strategies remain efficient: adapter-based tuning works well below 1k examples, while LoRA is effective for 1k–10k, as shown on $\text{BART}_{\text{large}}$ across text comparison tasks (Mao et al., 2022).