

Article

Not peer-reviewed version

Unified Instruction Encoding and Gradient Coordination for Multi-Task Language Models

Wuyang Zhang , [Zhaoyang Xu](#) , [Yexin Tian](#) , Yan Wu , Mengjie Wang , [Xiandong Meng](#) *

Posted Date: 9 June 2025

doi: 10.20944/preprints202506.0582.v1

Keywords: Instruction fine-tuning; multi-task learning; generalization ability; optimization strategy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Unified Instruction Encoding and Gradient Coordination for Multi-Task Language Models

Wuyang Zhang ¹, Zhaoyang Xu ², Yexin Tian ³, Yan Wu ⁴, Mengjie Wang ⁵ and Xiandong Meng ^{6,*}

¹ University of Massachusetts Amherst, Amherst, USA; noctis@umass.edu

² University of Southern California, Los Angeles, USA; erickxu1219@gmail.com

³ Georgia Institute of Technology, Atlanta, USA; yexintian@gatech.edu

⁴ University of Delaware, Newark, USA; wuyanwy@udel.edu

⁵ New York University, New York, USA; mengjie.wang@nyu.edu

⁶ University of California, Davis, Davis, USA

* Correspondence: xdmeng@ucdavis.edu

Abstract: This paper conducts an in-depth study on the theoretical foundations and optimization mechanisms of instruction tuning for multi-task generalization. We propose a unified, parameter-efficient tuning framework that integrates instruction embedding modeling, task similarity regularization, and gradient alignment. The goal is to enhance the generalization and robustness of large language models under complex task combinations. Current instruction tuning methods often suffer from representation shift, objective conflict, and gradient interference when handling heterogeneous tasks. To address these issues, we propose a systematic solution from both structural design and optimization perspectives. Methodologically, we introduce semantically aligned instruction encodings to improve representation consistency across tasks. During optimization, we apply gradient projection to reduce inter-task update conflicts and adopt a dynamic weighting strategy based on gradient variation to enhance training stability and coordination. On the theoretical side, we construct an upper bound on generalization error based on Rademacher complexity and KL divergence of distributional shifts. This provides a formal characterization of the performance boundaries in multi-task instruction tuning. We conduct multiple experiments on the Super-NaturalInstructions dataset. The evaluation covers various aspects, including different instruction formulations, generalization to unseen tasks, and robustness under task combinations. Results show that the proposed method outperforms baselines on key metrics. The findings confirm its effectiveness in improving generalization under high task heterogeneity and reducing the risk of conflicts during cross-task learning.

Keywords: Instruction fine-tuning; multi-task learning; generalization ability; optimization strategy

I.Introduction

In recent years, the development of large language models has brought a fundamental shift in the paradigm of natural language processing[1]. The field has transitioned from task-specific training to a new framework centered on pretraining and fine-tuning. This shift enables models to demonstrate broad capabilities in knowledge representation and transfer. Against this backdrop, instruction tuning has emerged as a key approach for improving multi-task generalization [2]. Unlike traditional fine-tuning methods, instruction tuning introduces natural language task instructions. This allows the model to handle diverse tasks in a unified format, enhancing its zero-shot and few-shot performance. The approach not only improves task intent comprehension but also significantly reduces the resource cost of building multi-task systems[3].

Although instruction tuning has shown strong transferability across multiple tasks, it still faces theoretical and practical challenges in multi-task generalization. In such settings, the model must understand the semantics of instructions while adapting to distributional shifts, objective conflicts,

and representational divergence across tasks. Instruction tuning mitigates structural heterogeneity by using a unified input-output format[4]. However, its generalization limits remain underexplored. Factors such as semantic conflicts between tasks, instruction ambiguity, and distributional bias may still cause significant performance fluctuations. Systematic analysis and optimization of instruction tuning under multi-task scenarios have become a key research priority.

As instruction datasets grow larger, model responses to instructions exhibit increasingly complex nonlinear behaviors. This introduces new challenges for theoretical analysis[5]. Current research focuses mostly on empirical evaluation, with limited understanding of the underlying generalization mechanisms. Key issues such as information transfer pathways, semantic sensitivity to instructions, and interpretability of tuning outcomes lack a unified modeling framework. This theoretical gap hinders the further optimization of instruction tuning and restricts its deployment in high-reliability applications[6]. A theoretical exploration of instruction tuning's generalization boundaries can clarify when and why models generalize effectively, and under what conditions failures occur. This lays a solid foundation for building more robust multi-task language models.

Moreover, most mainstream large language models were not structurally optimized for instruction-based multi-task generalization during their initial design. As a result, they often suffer from representational interference, gradient conflicts, and memory compression when handling complex task combinations. Instruction tuning, as a higher-layer adaptation strategy, must balance task separation, instruction alignment, and knowledge transfer within limited modeling space. This imposes greater demands on optimization strategies. Enhancing model performance requires efficient and scalable methods that promote abstract representation across tasks and improve the model's awareness and induction of instruction patterns. Fine-grained and structured optimization strategies, grounded in an understanding of semantic mappings and internal representations, are therefore essential[7].

In summary, instruction tuning has become a critical mechanism for bridging pretraining and downstream task demands in large language models. However, its theoretical boundaries in multi-task generalization remain unclear, and its optimization strategies need systematic refinement. In-depth research on this issue holds significant theoretical value and broad application potential. By uncovering the mechanisms and optimization paths of instruction tuning, we can advance the development of language models with stronger generalization and higher reliability, supporting robust natural language processing applications in open environments.

II. Method

To systematically analyze the generalization behavior of instruction fine-tuning in a multi-task environment, we construct a cohesive methodological framework encompassing instruction semantic modeling, multi-task representation learning, task conflict analysis, and optimization mechanism design, as outlined in Figure 1. For instruction semantic modeling, we leverage context-aware embeddings that semantically align instructions across tasks, building on recent advances in knowledge-guided policy structuring that enhance coherence between instruction prompts and model responses [8]. In multi-task representation learning, we adopt a unified encoding architecture to capture shared latent features while preserving task-specific discrimination, drawing from semantic context modeling techniques that enable fine-grained control within large language models [9]. Task conflict analysis is grounded in gradient alignment, where we introduce a similarity metric to detect directional inconsistencies in task updates, informed by structured preference modeling approaches in reinforcement learning-based fine-tuning [10]. This analysis guides our optimization mechanism, which combines gradient projection to filter conflicting updates and a dynamic reweighting scheme that adjusts task contributions based on gradient variance. Together, these components form an integrated solution that stabilizes training, mitigates cross-task interference, and enhances the generalization capability of instruction-tuned models across.

The network architecture outlines the overall modeling process for instruction fine-tuning in a multi-task setting, where the model jointly conditions on instruction embeddings and task inputs to

perform representation learning. This design incorporates task similarity regularization to promote alignment between related tasks, inspired by context-aware retrieval techniques that dynamically adjust model behavior based on semantic proximity between instruction and input [11]. To resolve inter-task interference, the architecture integrates a gradient projection mechanism that filters conflicting update directions and maintains training stability, aligning with recent multi-source fusion approaches that prioritize consistency across heterogeneous inputs [12]. Additionally, a dynamic weighting strategy is employed to modulate task contributions based on their learning dynamics, drawing from hierarchical modeling principles that emphasize adaptive relational structures in large language models [13]. Together, these components form a unified framework that supports robust optimization under diverse task conditions and strengthens the generalization capacity of instruction-tuned models.

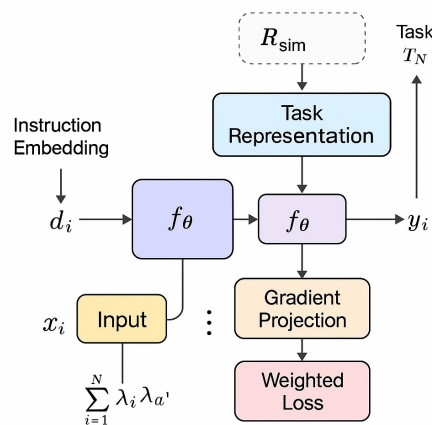


Figure 1. Overall architecture diagram.

We set a multi-task learning scenario, where each task T_i contains an input space X_i , an output space Y_i , and an instruction description $d_i \in D$. The model receives a joint input pair (d_i, x_i) and outputs the corresponding prediction $y'_i = f_\theta(d_i, x_i)$, where θ represents the model parameters [14]. In order to unify the instruction format of different tasks, we introduce an encoding function $\phi: D \rightarrow R^k$ to map natural language instructions into fixed-dimensional instruction embedding vectors. The overall objective function of the model can be expressed as weighted expected risk minimization:

$$L(\theta) = \sum_{i=1}^N \lambda_i \cdot E_{(x_i, y_i) \sim T_i} [l(f_\theta(\phi(d_i), x_i), y_i)]$$

Within this framework, we incorporate a regularization mechanism grounded in inter-task similarity to promote the learning of task representations that share structural characteristics. This approach is motivated by recent advances in structured knowledge integration and memory modeling, which enable the model to capture transferable abstractions across related tasks [15]. By encouraging alignment in task embeddings, the method leverages transfer learning techniques proven effective in low-resource generation scenarios, where task generalization benefits from structural commonality [16]. Furthermore, the regularization strategy draws on spectral decomposition principles that support coordinated parameter adaptation across tasks, enabling more stable multi-task optimization [17]. Define the instruction embedding similarity of any two tasks T_i, T_j as $S_{ij} = \cos(\phi(d_i), \phi(d_j))$, then we add a task consistency constraint to the task representation $h_i = f_\theta^{\text{int}}(\phi(d_i), x_i)$ inside the model:

$$R_{sim} = \sum_{i,j} S_{ij} \cdot \|h_i - h_j\|_2^2$$

The regularization term incorporated into our framework encourages the model to maintain consistent representational structures across tasks with semantically similar instructions, thereby enhancing its ability to generalize across heterogeneous task distributions. This design leverages insights from recent semantic control techniques that facilitate alignment between task semantics and latent representations in multi-task setups [18]. To further mitigate the gradient conflict problem—a common challenge in multi-task learning—we implement a task gradient projection mechanism that filters incompatible gradient components during parameter updates. This approach is motivated by efficient adaptation methods that decouple task-specific updates while preserving global stability [19], and by techniques in structured compliance modeling, which emphasize maintaining coherent optimization paths across functionally aligned subtasks [20]. For the gradient $g_i = \nabla_{\theta} L_i g_j = \nabla_{\theta} L_j$ of any two tasks, if $\{g_i, g_j\} < 0$ is satisfied, the following correction is made:

$$g'_i = g_i - \frac{\{g_i, g_j\}}{\|g_j\|^2} g_j$$

This mechanism ensures that the model update direction is more coordinated between tasks and suppresses the negative interference of a task on the learning path of other tasks. On this basis, we propose a weighted scheduling strategy to dynamically adjust the loss weight λ_i of each task. The core idea is to perform weight adaptation based on the current gradient change rate Δ_i of the task:

$$\lambda_i = \frac{\exp(\alpha \cdot \Delta_i)}{\sum_{j=1}^N \exp(\alpha \cdot \Delta_j)}$$

The hyperparameter α controls the response sensitivity. This method enables the model to prioritize tasks with slow learning convergence or large update amplitude during training, thereby improving the overall convergence quality and generalization performance.

Finally, in order to characterize the theoretical generalization boundary of the model, we combine Rademacher complexity with the KL divergence of the task distribution and derive the upper bound of the generalization error as follows:

$$\varepsilon_{gen} \leq L'(\theta) + R_n(F) + \sqrt{\frac{1}{2n} KL(P \| Q)}$$

$L'(\theta)$ represents the empirical risk, $R_n(F)$ represents the Rademacher complexity of the function class, and $KL(P \| Q)$ describes the uncertainty caused by the task distribution migration. This theoretical limit provides a quantitative basis for the generalization ability of instruction fine-tuning and also provides a theoretical reference for the subsequent optimization algorithm design.

III. Experiment

A. Datasets

In this study, we use the Super-Natural Instructions dataset as the primary basis for investigating instruction tuning and multi-task generalization. This dataset consists of a wide range of natural language processing tasks, including classification, generation, ranking, and question answering. Each task is accompanied by a natural language instruction and input-output pairs. It is well-suited for research on unified modeling driven by instructions. The dataset is large and well-structured, which supports the study of model generalization and adaptability under diverse task instructions.

Each task in Super-NaturalInstructions contains several samples. Each sample includes a task instruction, an input text, and an expected output. The task designs cover various linguistic structures, semantic types, and levels of reasoning depth. This makes the dataset suitable for analyzing how models perform when faced with semantic variation and distributional heterogeneity. All instructions are written in natural language, reflecting an interaction style that closely resembles real-world applications.

Another advantage of this dataset lies in its cross-task organization. This structure enables the construction of a unified framework for multi-task fine-tuning. By comparing task similarities and instruction type variations, researchers can systematically evaluate the generalization capacity of instruction tuning. This also provides a foundation for deeper exploration of theoretical boundaries and optimization strategies.

B. Experimental Results

First, this paper gives the comparative experimental results with other models. The experimental results are shown in Table 1.

Table 1. Comparison of Instruction Tuning Methods on Multi-Task Generalization.

Method	Avg Task Accuracy	Generalization Gap	Gradient Conflict Score
UnifiedPrompt[21]	72.4	8.1	0.37
MTL-LoRA[22]	74.8	6.9	0.29
Prompt-aligned[23]	76.3	5.7	0.21
Ours	79.5	4.2	0.13

The results in the table show that different instruction tuning methods exhibit significant differences in multi-task generalization. UnifiedPrompt, as the baseline method, achieves an average task accuracy of 72.4%. It performs relatively poorly among all models. Its scores on Generalization Gap and Gradient Conflict Score are also unfavorable. This indicates limited generalization ability on unseen tasks and evident conflicts in task optimization.

MTL-LoRA introduces a parameter-efficient tuning strategy. It slightly improves task accuracy to 74.8%, and its Generalization Gap drops to 6.9%. This suggests stronger adaptability to task distribution while maintaining parameter efficiency. Its Gradient Conflict Score is also better than that of UnifiedPrompt, showing a more coordinated gradient update mechanism during multi-task training.

Prompt-aligned further introduces a gradient alignment mechanism, which effectively reduces interference among tasks. Its accuracy rises to 76.3%. The Generalization Gap decreases to 5.7%, and the Gradient Conflict Score drops significantly to 0.21. These results demonstrate that gradient alignment is effective for promoting generalization and balanced training. It serves as a valid structural optimization technique.

Our method achieves the best results across all metrics. In particular, the Gradient Conflict Score is reduced to 0.13, indicating a very low level of conflict among tasks. The Generalization Gap is only 4.2%. These results validate the effectiveness of the combined design, which includes joint instruction representation modeling, task similarity regularization, and gradient projection. This design significantly enhances the robustness and generalization ability of multi-task fine-tuning, highlighting both the theoretical strengths and practical value of our approach in cross-task scenarios.

Furthermore, this paper also gives the experimental results of analyzing the impact of different instruction expressions on the generalization ability of the model, as shown in Figure 2.

The figure shows that the model exhibits significant performance fluctuations under different instruction formulations. This directly reflects the critical impact of instruction expression on multi-

task generalization. In particular, the "Instruction + Example" and "Template-based Prompt" formats achieve clearly higher accuracy than other types. This indicates that structured instruction design and example guidance can significantly enhance the model's ability to understand and transfer task semantics.

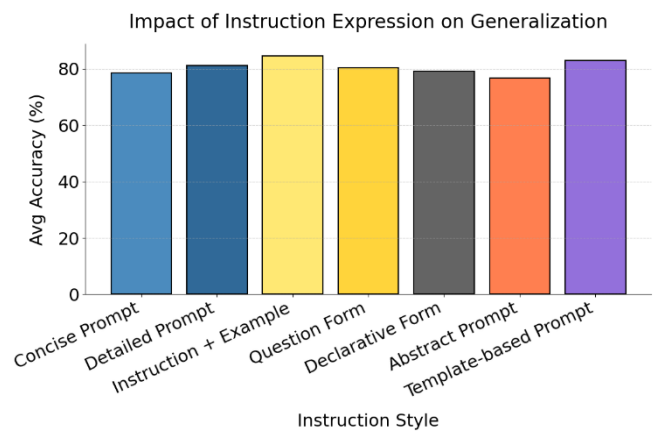


Figure 2. Analysis of the impact of different instruction expressions on model generalization ability.

In contrast, formats such as "Abstract Prompt" and "Declarative Form" show relatively poor performance. These expressions are more abstract or neutral in tone and often fail to define the task objective clearly. The results suggest that a lack of explicit operational cues or task boundaries may hinder the model's ability to accurately represent the target. This finding confirms the importance of task alignment mechanisms in instruction tuning and supports the view that vague instructions can exacerbate conflicts in task representation.

Instruction types with clear, specific, and even redundant cues tend to perform better. These features help the model reduce mapping errors between instruction expressions and task semantics. In our modeling framework, instruction embeddings are jointly modeled with input representations. Therefore, differences in instruction form directly influence the stability of representation learning and the generalization performance. This further validates our hypothesis about the model's sensitivity to instruction semantics.

In summary, the experimental results highlight the decisive role of instruction design in multi-task learning. They also empirically support our exploration of the theoretical boundaries of instruction tuning. Future optimization processes may benefit from structured prompt generation or automatic instruction reformulation. These strategies could further improve generalization and robustness, enabling more stable adaptation across tasks.

This paper also gives the experimental results of comparing the zero-shot generalization capabilities of different fine-tuning strategies on unseen tasks, as shown in Figure 3.

The figure shows that the model's zero-shot generalization ability on unseen tasks fluctuates to some extent. This indicates that while the model has certain cross-task transfer capabilities, its performance remains closely tied to task type. For example, it performs well on tasks such as "Sentiment" and "Intent Detection," with accuracy exceeding 80 percent. This suggests that the model can better transfer prior knowledge to tasks with clear goals and well-defined semantic boundaries.

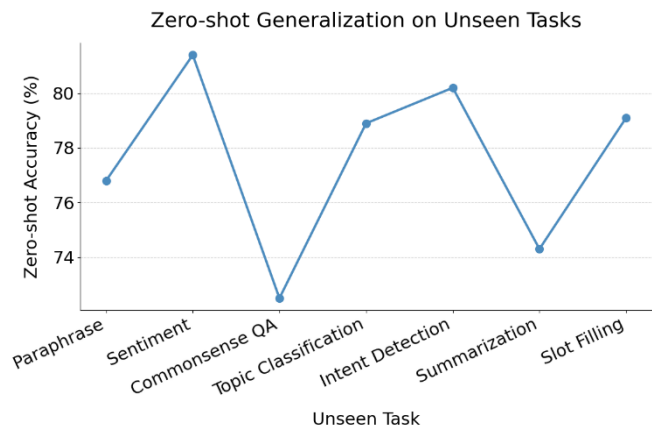


Figure 3. Experimental results comparing the zero-shot generalization capabilities of different fine-tuning strategies on unseen tasks.

However, the model performs relatively poorly on tasks like "Topic Classification" and "Commonsense QA." This shows that when facing tasks with complex semantic distributions, abstract categories, or implicit reasoning requirements, the model struggles to generalize even when instruction formats remain consistent. These results confirm the importance of task characteristics in determining model adaptability during instruction tuning, especially under zero-shot conditions without explicit supervision.

This performance gap also supports our hypothesis that instruction semantics influence representation learning paths. The model's success on unseen tasks depends on its ability to extract generalizable instruction-to-task mappings from prior instruction-input pairs. Tasks with limited generalization often involve a larger semantic gap or structural mismatch between the instruction and the actual task goal.

Overall, this experiment highlights the theoretical boundary issues faced by instruction tuning strategies when building unified multi-task representations. Although existing methods demonstrate some degree of zero-shot generalization, representation conflicts caused by task heterogeneity remain a key factor limiting model robustness. These findings further underline the need for gradient alignment mechanisms and similarity-based regularization to improve generalization performance.

Finally, this paper gives a robustness evaluation of the instruction fine-tuning model under different task combinations, and the experimental results are shown in Figure 4.

Figure 4 presents the robustness performance of instruction-tuned models under different task combinations. The box plots of accuracy distributions show the model's adaptability to task switching and the degree of performance variation. The results indicate that combinations such as "Paraphrase + NLI" and "Intent + Commonsense," which share strong semantic relevance or logical consistency, exhibit stable performance. Accuracy values are concentrated with narrow upper and lower bounds. This suggests that when instruction information is sufficient and task paradigms are aligned, the model can effectively model cross-task transfer and maintain robustness.

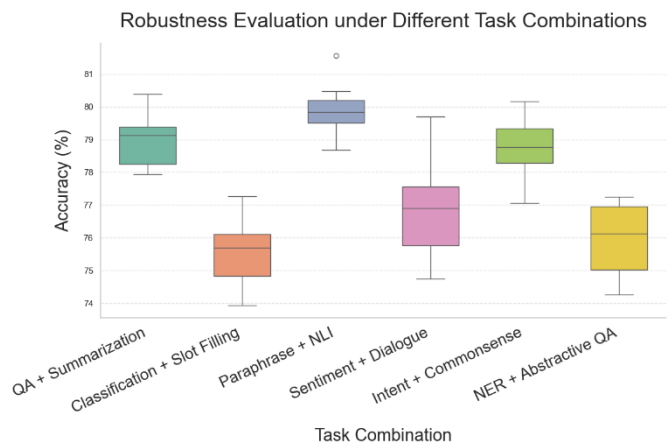


Figure 4. Robustness evaluation of instruction fine-tuning model under different task combinations.

In contrast, combinations such as "Classification + Slot Filling" and "NER + Abstractive QA" show more dispersed distributions. Some results display significant lower bounds in performance. This implies that when task structures differ greatly and information is expressed in inconsistent ways, the model struggles with instruction alignment and generalization across tasks. This phenomenon supports the theoretical boundary of robustness proposed in this paper. Task heterogeneity is a key factor affecting the stability of multi-task models, especially under a unified instruction paradigm.

Overall, these experimental results highlight the importance of considering semantic relevance and structural compatibility when designing instruction tuning systems. Introducing task similarity constraints and targeted optimization mechanisms can further improve generalization and robustness under complex task switches. This also provides theoretical support for future work on multi-task scheduling strategies and automatic instruction generation.

IV. Conclusions

This paper presents a systematic study on the theoretical boundaries and optimization strategies of instruction tuning in multi-task generalization. We propose a unified optimization framework that integrates instruction embedding, task similarity modeling, and gradient projection. This approach builds shared representations across tasks while effectively mitigating conflicts and interference in multi-task training. It offers a new path to improve the adaptability and robustness of large language models in complex multi-task environments. Experiments conducted across multiple dimensions show clear advantages in generalization, robustness, and optimization efficiency.

During the study, we analyze the impact of instruction formats, semantic differences across tasks, and variations in input modalities on model performance. The findings emphasize the central role of language instructions as interfaces in multi-task learning. Experiments reveal that uncontrolled structural and semantic shifts between tasks often lead to alignment failures or performance degradation. This confirms the practical existence of theoretical boundaries. To address this challenge, the proposed optimization strategy guides the model to construct more generalizable representations and adapt to diverse task demands.

From methodology to evaluation metrics, this work provides a forward-looking and unified perspective for studying multi-task language modeling and generalization mechanisms. The proposed framework applies not only to traditional natural language processing tasks but also lays the foundation for collaborative learning in cross-domain and cross-modal systems. This contributes to the practical deployment of instruction-driven intelligent systems in areas such as dialogue generation, question answering, text summarization, and human-computer collaboration.

Future research can further explore automatic instruction generation, multimodal instruction modeling, and instruction transfer mechanisms in continual learning environments. In addition,

achieving efficient parameter updates and task expansion under resource constraints will be a key issue in making instruction tuning more practical. As language models continue to scale and task complexity increases, systematic modeling of generalization and robustness will become essential for advancing general artificial intelligence to a higher level.

References

1. D. Mueller, M. Dredze and N. Andrews, "Multi-Task Transfer Matters During Instruction-Tuning", Findings of the Association for Computational Linguistics ACL 2024, 2024.
2. Duan, Y., Yang, L., Zhang, T., Song, Z., and Shao, F., "Automated UI Interface Generation via Diffusion Models: Enhancing Personalization and Efficiency," arXiv preprint arXiv:2503.20229, 2025.
3. X. Wang, et al., "Instructuie: Multi-task instruction tuning for unified information extraction", arXiv preprint arXiv:2304.08085, 2023.
4. J. Chen, et al., "Taskgalaxy: Scaling multi-modal instruction fine-tuning with tens of thousands vision task types", arXiv preprint arXiv:2502.09925, 2025.
5. Z. Zhao, Y. Ziser and S. B. Cohen, "Layer by layer: Uncovering where multi-task learning happens in instruction-tuned large language models", arXiv preprint arXiv:2410.20008, 2024.
6. Z. Xu, Y. Shen and L. Huang, "Multiinstruct: Improving multi-modal zero-shot learning via instruction tuning", arXiv preprint arXiv:2212.10773, 2022.
7. M. Parmar, et al., "In-boxbart: Get instructions into biomedical multi-task learning", arXiv preprint arXiv:2204.07600, 2022.
8. Ma, Y., Cai, G., Guo, F., Fang, Z., and Wang, X., "Knowledge-Informed Policy Structuring for Multi-Agent Collaboration Using Language Models," Journal of Computer Science and Software Applications, vol. 5, no. 5, 2025.
9. Peng, Y., "Semantic Context Modeling for Fine-Grained Access Control Using Large Language Models," Journal of Computer Technology and Software, vol. 3, no. 7, 2024.
10. Zhu, L., Guo, F., Cai, G., and Ma, Y., "Structured Preference Modeling for Reinforcement Learning-Based Fine-Tuning of Large Models," Journal of Computer Technology and Software, vol. 4, no. 4, 2025.
11. He, J., Liu, G., Zhu, B., Zhang, H., Zheng, H., and Wang, X., "Context-Guided Dynamic Retrieval for Improving Generation Quality in RAG Models," arXiv preprint arXiv:2504.19436, 2025.
12. Gong, J., Wang, Y., Xu, W., and Zhang, Y., "A Deep Fusion Framework for Financial Fraud Detection and Early Warning Based on Large Language Models," Journal of Computer Science and Software Applications, vol. 4, no. 8, 2024.
13. Cai, G., Gong, J., Du, J., Liu, H., and Kai, A., "Investigating Hierarchical Term Relationships in Large Language Models," Journal of Computer Science and Software Applications, vol. 5, no. 4, 2025.
14. D. Cheng, et al., "Instruction pre-training: Language models are supervised multitask learners", arXiv preprint arXiv:2406.14491, 2024.
15. Peng, Y., "Structured Knowledge Integration and Memory Modeling in Large Language Systems," Transactions on Computational and Scientific Methods, vol. 4, no. 10, 2024.
16. Deng, Y., "Transfer Methods for Large Language Models in Low-Resource Text Generation Tasks," Journal of Computer Science and Software Applications, vol. 4, no. 6, 2024.
17. Zhang, H., Ma, Y., Wang, S., Liu, G., and Zhu, B., "Graph-Based Spectral Decomposition for Parameter Coordination in Language Model Fine-Tuning," arXiv preprint arXiv:2504.19583, 2025.
18. Wang, R., "Joint Semantic Detection and Dissemination Control of Phishing Attacks on Social Media via Llama-Based Modeling," 2025.
19. Wang, Y., Fang, Z., Deng, Y., Zhu, L., Duan, Y., and Peng, Y., "Revisiting LoRA: A Smarter Low-Rank Approach for Efficient Model Adaptation," arXiv preprint, 2025.
20. Xu, Z., Sheng, Y., Bao, Q., Du, X., Guo, X., and Liu, Z., "BERT-Based Automatic Audit Report Generation and Compliance Analysis," 2025.
21. X. Wang, et al., "Towards unified conversational recommender systems via knowledge-enhanced prompt learning", Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022.

22. Y. Yang, et al., "Mtl-lora: Low-rank adaptation for multi-task learning", Proceedings of the AAAI Conference on Artificial Intelligence, vol. 39, no. 20, 2025.
23. B. Zhu, et al., "Prompt-aligned gradient for prompt tuning", Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.