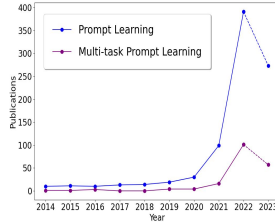


Motivation

> Rapid Progress on Prompt-based Learning Methods due to Large Language Models (LLMs)

> Recent methods focus on Multi-tasking capabilities.

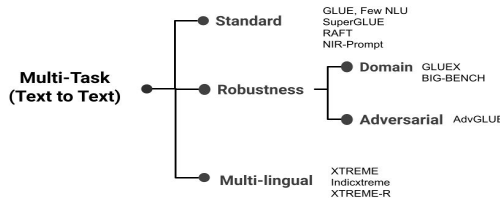
> Lack of Systematic Review.



Goals:

- > Organize & Compare the literature on Multi-Task Prompt Based Learning.
- > Identify Emerging & Future Directions.

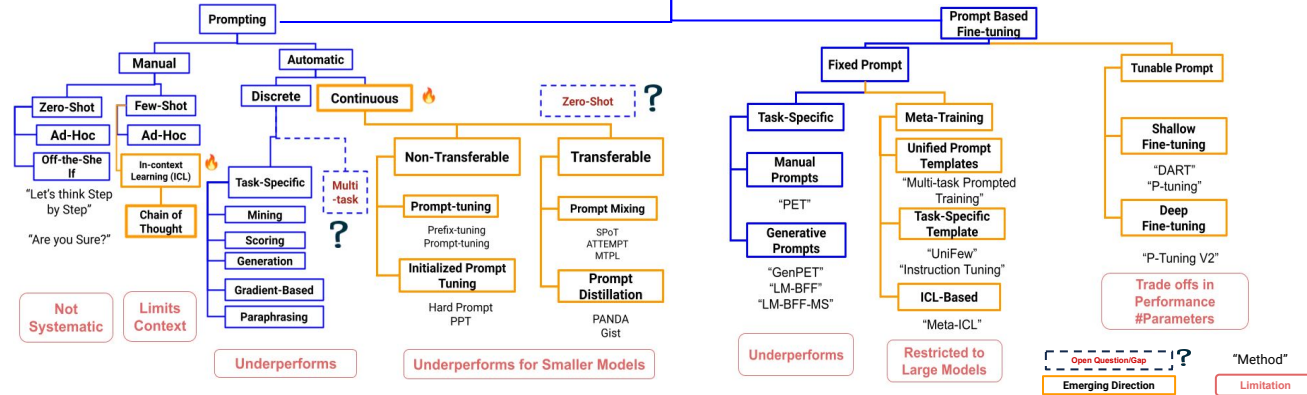
Benchmark Datasets



Limitations

- > **Misleading Metrics**: MCQ Evaluation, Non-Linear Metrics?
- > **Negative Task Transfer**: Which tasks should be combined to form dataset?
- > **Presence in Pre-training Dataset**: Memorization in LLMs

Multi-task Prompt Based Learning



Emergent Tasks?

> **Definition**: "An ability is emergent if it is not present in smaller models but is present in larger models." [1]

> **Scaling Factors**: Compute for Training, Number of Parameters, Dataset Size

> **Emergent Tasks**: Truthfulness, Toxicity classification, BIG-BENCH

Arguments against Emergence

- > **Metrics**: Using non-linear methods of evaluation metrics (accuracy) gives a "mirage" of emergence, Linear metrics (per token probability) show no evidence of emergence, [2]
- > **Training Data**: No clarity if same or similar tasks in pre-training data, different training data for models

Future Work

- > **Parameter-Efficient Methods (PEFT)**: [3] Add more Parameter Efficient Methods
- > **Other Modalities, Multi-Modalities**: Explore Video, Audio, Images, Multi-modal prompt based methods [4,5]

Future Directions

- > **Continuous Prompts**:
 - Applications: Using Prompts to bounds outputs, Symbolic Operations
 - Architecture Agnostic Prompts: Cross Architecture Prompt Use
 - Knowledge Transfer: Zero-Shot Way to transfer knowledge
 - Understanding Tradeoffs of Parameter Efficient Methods
- > **In-context Learning**:
 - Scaling: Make it Insensitive to Order of Examples
 - Investigation: How does it work + Which Tasks Perform Well?
- > **Benchmark Datasets**: Use Linear Metrics?, Better ways to
- > **Multi-Modal Applications**: E.g Unified-IO, Multi-Modal Benchmark Datasets
- > **Memorization, Biases, Negative Task Transfer**: Investigation in Large Language Models [5]

References

- [1] Wei, Jason, et al. "Emergent abilities of large language models." *arXiv preprint arXiv:2206.07682* (2022).
- [2] Schaeffer et al. "Are Emergent Abilities of Large Language Models a Mirage?"
- [3] Ding, Ning, et al. "Delta tuning: A comprehensive study of parameter efficient methods for pre-trained language models." *arXiv preprint arXiv:2203.06904* (2022).
- [4] Khatik, Muhammad Uzair, et al. "Maple: Multi-modal prompt learning." *arXiv preprint arXiv:2210.03117* (2022).
- [5] Zhao, Jinming, et al. "Memobert: Pre-training model with prompt-based learning for multimodal emotion recognition." *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022.
- [6] Akyurek, Afra Feyza, et al. "On measuring social biases in prompt-based multi-task learning." *arXiv preprint arXiv:2205.11605* (2022).