Multitask Prompted Training Enables Zero Shot Task Generalization

MULTITASK PROMPTED TRAINING ENABLES ZERO-SHOT TASK GENERALIZATION

Victor Sanh* Hugging Face
Albert Webson* Brown University
Colin Raffel* Hugging Face
Stephen H. Bach* Brown University

Lintang Sutawika BigScience
Zaid Alyafeai KFUPM
Antoine Chaffin IRISA, IMATAG
Arnaud Stiegler Hyperscience
Teven Le Scao Hugging Face

Arun Raja I²R, A*STAR
Manan Dey SAP
M Saiful Bari NTU
Canwen Xu UCSD/HF
Urmish Thakker SambaNova Systems

Shanya Sharma Walmart Labs
Eliza Szczechla BigScience
Taewoon Kim VU Amsterdam
Gunjan Chhablani BigScience
Nihal V. Nayak Brown

Debajyoti Datta University of Virginia
Jonathan Chang ASUS
Mike Tian-Jian Jiang ZEALS
Han Wang NYU
Matteo Manica IBM Research

Sheng Shen UC Berkeley
Zheng-Xin Yong Brown University
Harshit Pandey BigScience
Michael McKenna Parity
Rachel Bawden Inria, France

Thomas Wang Inria, France
Trishala Neeraj BigScience
Jos Rozen BigScience
Abheesht Sharma BITS Pilani, India
Andrea Santilli Sapienza

Thibault Fevry BigScience
Jason Alan Fries Stanford
Ryan Teehan CRA
Stella Biderman EleutherAI
Leo Gao EleutherAI
Tali Bers Brown

Thomas Wolf Hugging Face
Alexander M. Rush Hugging Face

ABSTRACT

Large language models have recently been shown to attain reasonable zero-shot generalization on a diverse set of tasks (Brown et al., 2020). It has been hypothesized that this is a consequence of implicit multitask learning in language model training (Radford et al., 2019). Can zero-shot generalization instead be directly induced by *explicit* multitask learning? To test this question at scale, we develop a system for easily mapping general natural language tasks into a human-readable prompted form. We convert a large set of supervised datasets, each with multiple prompts using varying natural language. These prompted datasets allow for benchmarking the ability of a model to perform completely unseen tasks specified in natural language. We fine-tune a pretrained encoder-decoder model (Raffel et al., 2020; Lester et al., 2021) on this multitask mixture covering a wide variety of tasks. The model attains strong zero-shot performance on several standard datasets, often outperforming models up to 16× its size. Further, our approach attains strong performance on a subset of tasks from the BIG-Bench benchmark, outperforming models up to 6× its size. All prompts and trained models are available at `github.com/bigscience-workshop/promptsource/` and `huggingface.co/bigscience/T0pp`.

1 INTRODUCTION

Recent work has shown that large language models exhibit the ability to perform reasonable zero-shot generalization to new tasks (Brown et al., 2020; Kim et al., 2021). Despite only being trained on language modeling objectives, these models can perform relatively well at new tasks that they have not been explicitly trained to perform, for instance answering a question on a passage or performing

*Equal contribution. Full list of individual contributions detailed in Appendix A.

1

Multitask prompted training enables zero shot task generalization is a groundbreaking approach in the field of artificial intelligence and natural language processing. This innovative methodology has emerged as a response to the growing demand for AI systems capable of seamlessly adapting to new tasks without the need for extensive retraining. By leveraging the principles of multitask learning and prompt engineering, researchers are exploring avenues to enhance generalization capabilities in AI models, allowing them to perform effectively in previously unseen scenarios.

Understanding Multitask Learning

Multitask learning is a machine learning paradigm where multiple learning

tasks are solved simultaneously, sharing representations and knowledge across tasks. This approach capitalizes on the idea that many tasks may share common features, thereby allowing a model to generalize better across different tasks.

Key Benefits of Multitask Learning

1. Improved Generalization: By training on multiple related tasks, models learn to capture shared structures, which can improve performance on unseen tasks.
2. Reduced Overfitting: Sharing parameters among tasks can reduce the risk of overfitting to any single task, as the model learns to generalize from diverse data.
3. Efficient Use of Data: Multitask learning enables the model to utilize data from various tasks, leading to better performance when data for a specific task is scarce.

Introduction to Prompting in AI

Prompting, particularly in the context of language models, refers to the technique of providing a model with specific cues or instructions to elicit desired outputs. This method has gained popularity due to its effectiveness in guiding models to understand and generate relevant responses based on the provided context.

Types of Prompts

- Task-specific Prompts: These prompts are tailored for particular tasks, helping the model focus on the required output.
- Instruction-based Prompts: These provide guidance on how to perform a task, often in the form of questions or directives.
- Few-shot Prompts: These include a few examples of the desired task, allowing the model to infer the pattern and apply it to new inputs.

Zero-Shot Learning and Generalization

Zero-shot learning refers to the ability of a model to perform a task that it has not been explicitly trained on. This capability is particularly valuable in real-world applications where the model may encounter novel tasks or situations.

How Zero-Shot Learning Works

Zero-shot learning relies on the model's capacity to leverage its existing knowledge and transfer learning from related tasks. This process typically involves:

1. Semantic Understanding: The model uses its training on related tasks to

form a semantic understanding of the new task.
2. Contextual Awareness: By interpreting prompts correctly, the model can adapt its responses to align with the requirements of the new task.
3. Knowledge Transfer: The model applies learned features and representations from similar tasks to tackle the new challenge.

The Role of Multitask Prompted Training

Multitask prompted training is a synthesis of multitask learning and prompt engineering. This approach involves training a model on multiple tasks simultaneously while utilizing prompts to direct its learning.

Key Components of Multitask Prompted Training

1. Diverse Task Selection: Choosing a variety of tasks that share underlying characteristics to enhance the model's learning capacity.
2. Effective Prompt Design: Crafting prompts that can effectively guide the model in understanding the nuances of different tasks.
3. Shared Representations: Developing a model architecture that encourages the sharing of knowledge across tasks, enhancing its ability to generalize.

Benefits of Multitask Prompted Training for Zero-Shot Generalization

The integration of multitask prompted training significantly enhances zero-shot generalization abilities in AI models. Here’s how:

1. Holistic Learning Approach

By exposing the model to various tasks during training, it builds a comprehensive understanding of language and context. This holistic learning enables the model to apply its knowledge effectively when faced with new tasks that share similarities with those it has learned.

2. Enhanced Prompting Strategies

Multitask prompted training allows for the development of more sophisticated prompting strategies. The model learns to interpret prompts not just in isolation but in the context of what it has learned across tasks. This capability is crucial for zero-shot scenarios, where the model must rely solely on its training to infer the correct actions or responses.

3. Robustness to Novelty

Training on multiple tasks helps the model become more robust to novel

inputs. It learns to adapt its responses based on various contexts and variations, making it less likely to falter when encountering unfamiliar tasks.

Challenges and Considerations

While multitask prompted training presents numerous advantages, it is not without its challenges.

1. Task Interference

Training on multiple tasks can sometimes lead to interference, where learning one task negatively impacts performance on another. Careful selection of tasks and balancing their complexity is essential to mitigate this issue.

2. Prompt Design Complexity

Creating effective prompts for diverse tasks can be complex. Poorly designed prompts may lead to confusion or misinterpretation, ultimately hindering the model’s performance.

3. Evaluation Difficulties

Evaluating the model's performance in zero-shot scenarios can be challenging, as traditional metrics may not adequately capture its generalization capabilities. Developing new evaluation frameworks is essential for assessing the effectiveness of multitask prompted training.

Future Directions

The field of multitask prompted training enabling zero-shot task generalization is still evolving. Future research may focus on:

1. Automated Prompt Generation: Developing systems that can automatically generate effective prompts based on the tasks at hand.
2. Dynamic Task Selection: Implementing methods for dynamically selecting tasks during training to optimize learning pathways.
3. Broader Application Domains: Exploring the application of multitask prompted training in various domains, including healthcare, finance, and autonomous systems.

Conclusion

In conclusion, the integration of multitask prompted training enables zero-shot task generalization represents a significant advancement in AI and natural language processing. By combining the strengths of multitask learning

and effective prompting strategies, models can achieve impressive generalization capabilities, allowing them to adapt to new tasks with minimal intervention. As research continues to progress in this area, we can expect to see even more sophisticated AI systems that are capable of tackling a wide array of challenges across diverse domains. The future of AI looks promising, with multitask prompted training paving the way for more intelligent and adaptable systems.

Frequently Asked Questions

What is multitask prompted training?

Multitask prompted training refers to a training approach where a model learns to perform multiple tasks simultaneously by leveraging shared prompts or contexts, enhancing its ability to generalize across different tasks.

How does multitask prompted training enable zero-shot task generalization?

By exposing the model to a variety of tasks during training, multitask prompted training allows the model to learn underlying patterns and relationships, enabling it to generalize to new, unseen tasks without the need for additional training.

What are the benefits of zero-shot task generalization in AI models?

Zero-shot task generalization allows AI models to adapt to new tasks efficiently without retraining, saving time and resources while also improving their flexibility and applicability in diverse scenarios.

Can you provide an example of multitask prompted training?

An example would be training a language model on tasks like sentiment analysis, text summarization, and question answering simultaneously using prompts that indicate the desired task, allowing it to handle any of these tasks in new contexts with minimal adjustments.

What types of tasks can benefit from multitask prompted training?

Tasks in natural language processing, computer vision, and even reinforcement learning can benefit from multitask prompted training, as it encourages models to share knowledge and improve performance across varied applications.

What challenges might arise with multitask prompted training?

Challenges include managing conflicting information from different tasks, ensuring balanced training across tasks to avoid overfitting on one, and the complexity of designing effective prompts that guide the model appropriately.

How does this training method compare to traditional single-task training?

Unlike traditional single-task training, which focuses on one specific task, multitask prompted training fosters broader generalization capabilities by allowing the model to learn from multiple tasks at once, enhancing its robustness and adaptability.

Multitask Prompted Training Enables Zero Shot Task Generalization

Responsibility in Gaming
As an evolution of traditional CSR, ESG adds critical accountability through transparent measurement, monitoring and reporting, which are fast becoming the …

AGA: Smashing ESG - GGB News
Apr 25, 2022 · When it comes to sustainability, diversity and responsible operations, the gaming industry is hitting its marks. The American Gaming …

The Gaming Industry's Commitment to ESG - American …
“This is just the beginning for gaming and ESG. Working with our members, the AGA will expand our industry’s positive impact, defining new ways to help gaming lead …

AGA publishes ESG in Gaming: Our Industry’s Commitments
Our members efforts are improving the planet, gaming communities and our employee and customer experience. “This is just a starting point for the AGA and …

Sustainability, ESG and safer gambling - IAGR - International ...
Jan 28, 2025 · United States In 2022, the American Gaming Association conducted an industry-wide assessment of its member organizations’ strategies and …

Hollókő - Wikipedia
Its name means "Raven-stone" in Hungarian. The village is located in Nógrád county, approximately 91.1 kilometres northeast from Budapest, the capital of Hungary. It lies in a valley of Cserhát Mountains, surrounded by low peaks. The natural environment is protected.

Old Village of Hollókő (2025) - All You Need to Know ... - Tripadvisor
Hollókő is a Paloc ethnographic village in Hungary and a UNESCO World Heritage Site.

Old Village of Hollókő and its surroundings
Hollókő in Nógrád County in Northern Hungary has been a UNESCO World Heritage site. This small village is sought out by tourists from all over the world, as this is currently the only Hungarian

settlement whose residents still use the built environment as …

Old Village of Hollókő and its Surroundings - UNESCO World …
The Old Village of Hollókő is a Palócz settlement located in the County of Nógrád in Northern Hungary, about 100 km north-east of Budapest. The Old Village, which has been deliberately preserved, is a living example of rural life before the agricultural revolution of the 20th century.

How to visit Holloko from Budapest in 2025 - Time Travel Turtle
May 18, 2025 · Hungary's World Heritage village, Holloko, is a fascinating insight into traditional life on the land, where you can visit a perfectly preserved community.

The unique Hungarian village: Hollókő - Info Budapest
Jul 6, 2020 · The small village in Nógrád County is the only settlement in Hungary that is on the UNESCO World Heritage List. It didn’t gain this honor in any way but through authentically presenting the untimely 20th-century village life.

Travel Back in Time in the Charming Village of Hollókő!
Hollókő is located in the northeast part of Hungary, about a 90-minute drive from Budapest. A UNESCO World Heritage Site since 1987, the village and its surroundings offer a unique insight into what life used to be like in rural Hungary before the 20th century.

Hollókö: UNESCO World Heritage Site Travel Guide
Hollókö is a Paló c ethnographic village, an ethnic minority group in north-east Hungary. The village developed mainly during the 18th and 19th centuries and has preserved its half-timbered architecture. The surroundings include strip-field farming, orchards, vineyards, meadows and …

Day Trip to Hollókő (Visit the Village, the Castle and a Secret Spot)
Set out on a journey to Hollókő and learn about the rich history and culture of Hungary. Explore the charming village, visit the historic castle, and uncover a hidden gem off the beaten path.

Old Village of Hollókő and its Surroundings, Hungary
Hollókő is a small village in Northern Hungary which was granted World Heritage status in 1987 as the country’s only living example of a perfectly preserved settlement from the last century,…