Transformer^2 : Self-adaptive LLMs

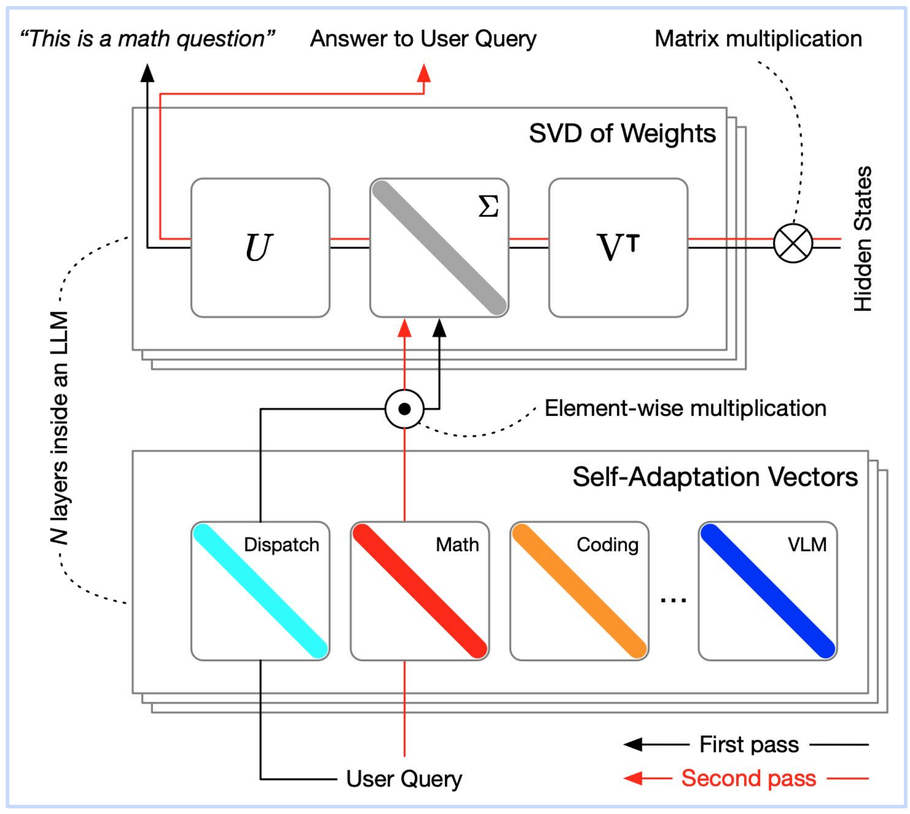
Qi Sun*, Edoardo Cetin*, Yujin Tang* *{qisun, edo, yujintang}@sakana.ai*
*Equal contribution
sakana.ai

Fine Tuning and Adaptive LLMs

- Self-adaptive LLMs would let models adjust to tasks in real-time.
- Traditional training tries to optimize for all tasks at once but is inefficient.
- Expert modules could help but face issues with large number parameters, overfitting, and composition.

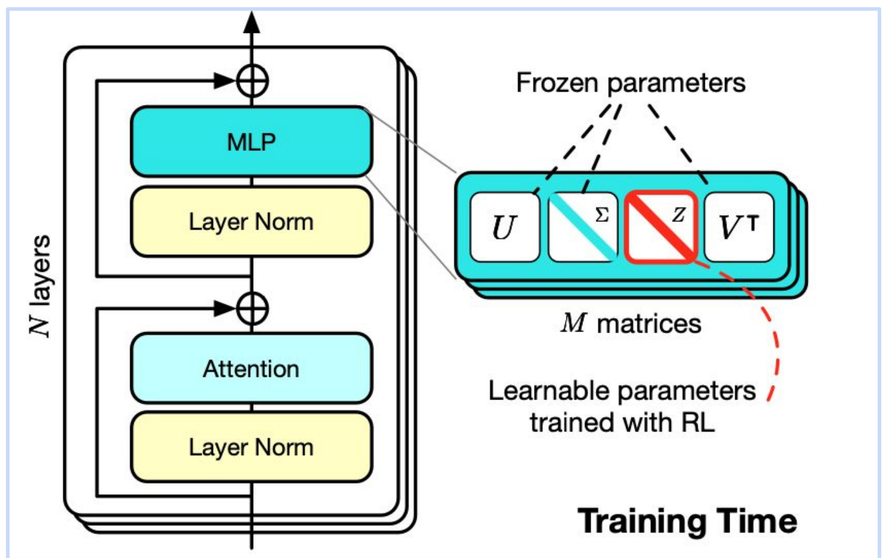
Transformer^2

A **self-adaptation** framework that adapts LLMs for unseen tasks in real-time.



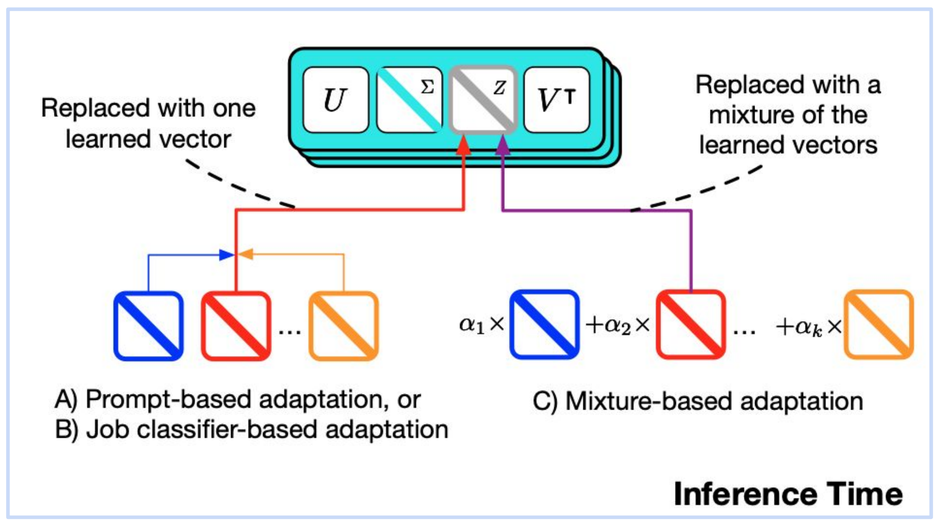
• Singular Value Finetuning

In training, we tune the scales of the **singular values** of the weight matrices to generate a set of "expert" vectors.



• Adaptive Inference

In inference, a **two-pass** process is adopted where the first applies the task- specific expert and the second generates the answer.

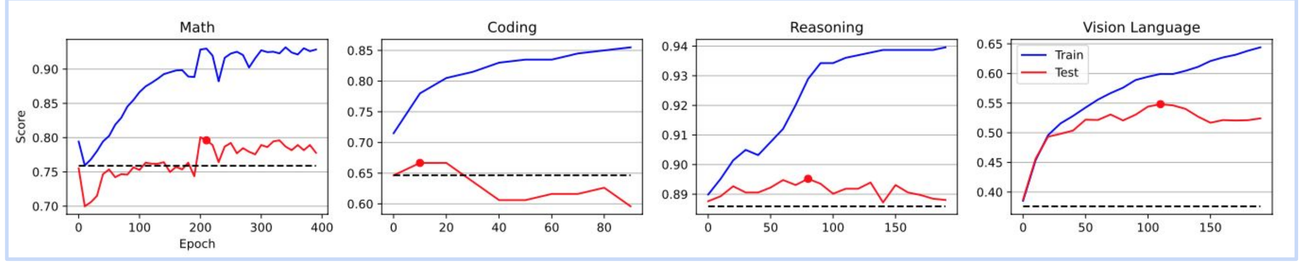


Training Experts Vectors

Expert vectors are trained through **reinforcement learning**, using **KL-divergence** as a regularization technique.

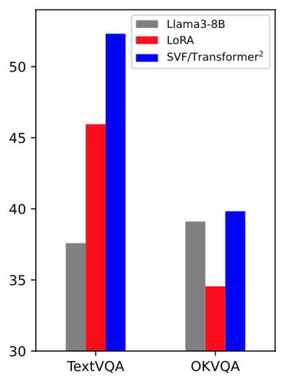
- Reduces dependency on specific training datasets
- Allows direct task-specific feedback to improve model performance

$$J(\theta_z) = \mathbb{E}\left[\log\left(\pi_{\theta_{W'}}(\hat{y}_i \mid x_i)\right) r(\hat{y}_i, y_i)\right] - \lambda D_{\mathrm{KL}}(\pi_{\theta_{W'}} \| \pi_{\theta_W})$$



Result 1 - Seen Tasks

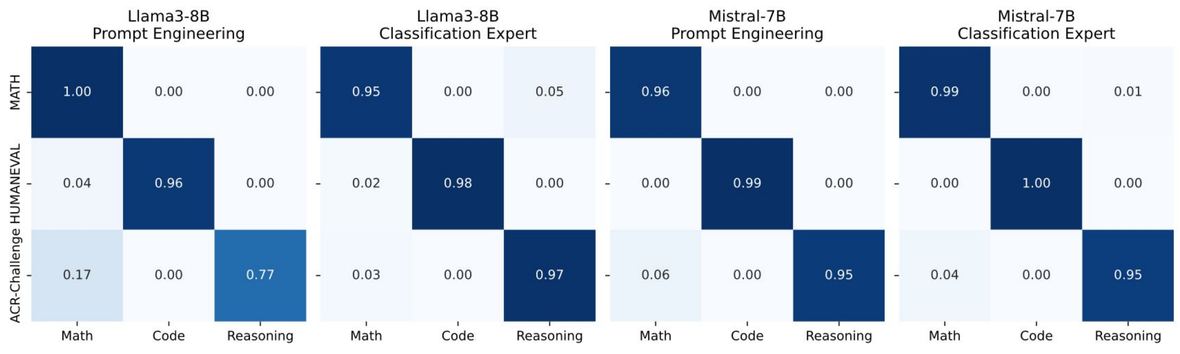
Method	GSM8K	MBPP-Pro	ARC-Easy
LLAMA3-8B-INSTRUCT	75.89 (1.00)	64.65 (1.00)	88.59 (1.00)
+ LoRA	77.18 (1.02)	**67.68** (**1.05**)	88.97 (1.00)
+ SVF (Ours)	**79.15** (**1.04**)	66.67 (1.03)	**89.56** (**1.01**)
MISTRAL-7B-INSTRUCT-V0.3	42.83 (1.00)	49.50 (1.00)	81.65 (1.00)
+ LoRA	44.66 (1.04)	51.52 (1.04)	81.19 (0.98)
+ SVF (Ours)	**49.74** (**1.16**)	**51.52** (**1.04**)	**85.14** (**1.04**)
LLAMA3-70B-INSTRUCT	85.29 (1.00)	**80.81** (**1.00**)	**89.10** (**1.00**)
+ LoRA	77.26 (0.91)	68.69 (0.85)	88.55 (0.99)
+ SVF (Ours)	**88.32** (**1.04**)	**80.81** (**1.00**)	88.47 (0.99)



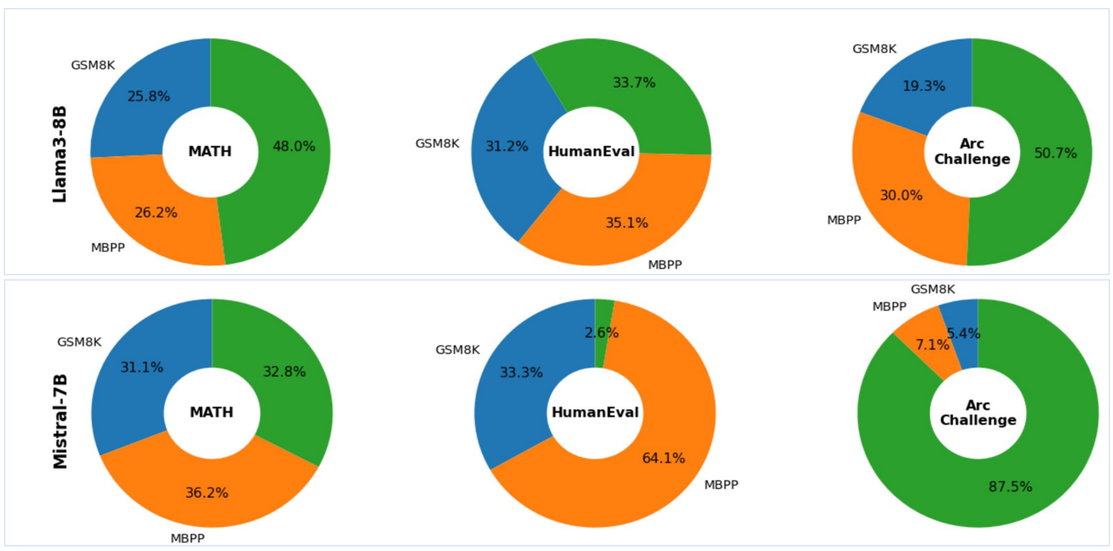
Result 2 - Unseen Tasks

Method	MATH	Humaneval	ARC-Challenge
LLAMA3-8B-INSTRUCT 3	24.54 (1.00)	60.98 (1.00)	80.63 (1.00)
+ LoRA	24.12 (0.98)	52.44 (0.86)	81.06 (1.01)
+ Transformer^2 (Prompt)	25.22 (1.03)	61.59 (1.01)	81.74 (1.01)
+ Transformer^2 (Cls-expert)	25.18 (1.03)	62.80 (1.03)	81.37 (1.01)
+ Transformer^2 (Few-shot)	**25.47** (**1.04**)	**62.99** (**1.03**)	**82.61** (**1.02**)
MISTRAL-7B-INSTRUCT-V0.3	13.02 (1.00)	43.29 (1.00)	71.76 (1.00)
+ LoRA	13.16 (1.01)	37.80 (0.87)	**75.77** (**1.06**)
+ Transformer^2 (Prompt)	11.86 (0.91)	43.90 (1.01)	72.35 (1.01)
+ Transformer^2 (Cls-expert)	11.60 (0.89)	43.90 (1.01)	74.83 (1.04)
+ Transformer^2 (Few-shot)	**13.39** (**1.03**)	**47.40** (**1.09**)	75.47 (1.05)
LLAMA3-70B-INSTRUCT	**40.64** (**1.00**)	78.66 (1.00)	87.63 (1.00)
+ LoRA	25.40 (0.62)	73.78 (0.94)	83.70 (0.96)
+ Transformer^2 (Prompt)	40.44 (1.00)	**79.88** (**1.02**)	**88.48** (**1.01**)

Analysis 1 - Job dispatching accuracy



Analysis 2 - Adaptation contribution



Analysis 3 - Cross-model compatibility

Method *SVF training task*	MATH *GSM8K*	Humaneval *MBPP-pro*	ARC-Challenge *ARC-Easy*
MISTRAL-7B-INSTRUCT-V0.3	**13.02** (**1.00**)	43.29 (1.00)	71.76 (1.00)
+ Llama SVF (ordered σ_i)	11.96 (0.92)	45.12 (1.04)	72.01 (1.00)
+ Llama SVF (shuffled σ_i)	10.52 (0.81)	40.24 (0.93)	70.82 (0.99)
+ Few-shot adaptation (cross-model)	12.65 (0.97)	**46.75** (**1.08**)	**75.64** (**1.05**)

Future Direction

- Explore model merging techniques to overcome the limitations of SVD experts' dependency on base models.
- Investigate ways to reduce computational costs when scaling to multiple specialized domains.
- Leverage advanced techniques to develop more powerful adaptation strategy.