EXPLORA: Efficient Exemplar Subset Selection for Complex Reasoning

Kiran Purohit, Venktesh V, Raghuram Devalla, Krishna Mohan Yerragorla,
Sourangshu Bhattacharya, Avishek Anand

Dept. of Computer Science & Engineering
IIT Kharagpur

In-Context Learning (ICL)

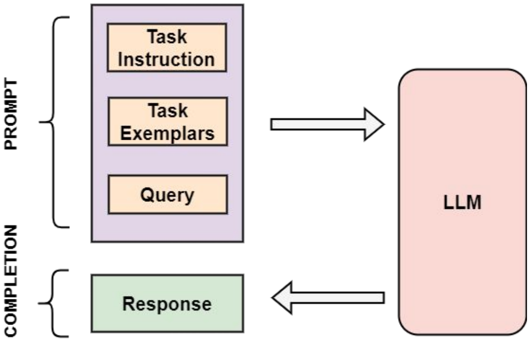
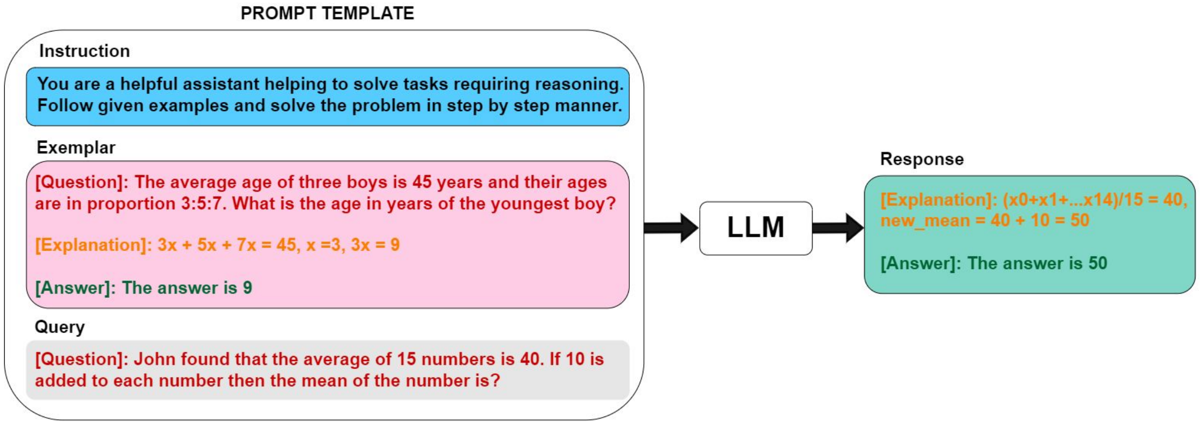


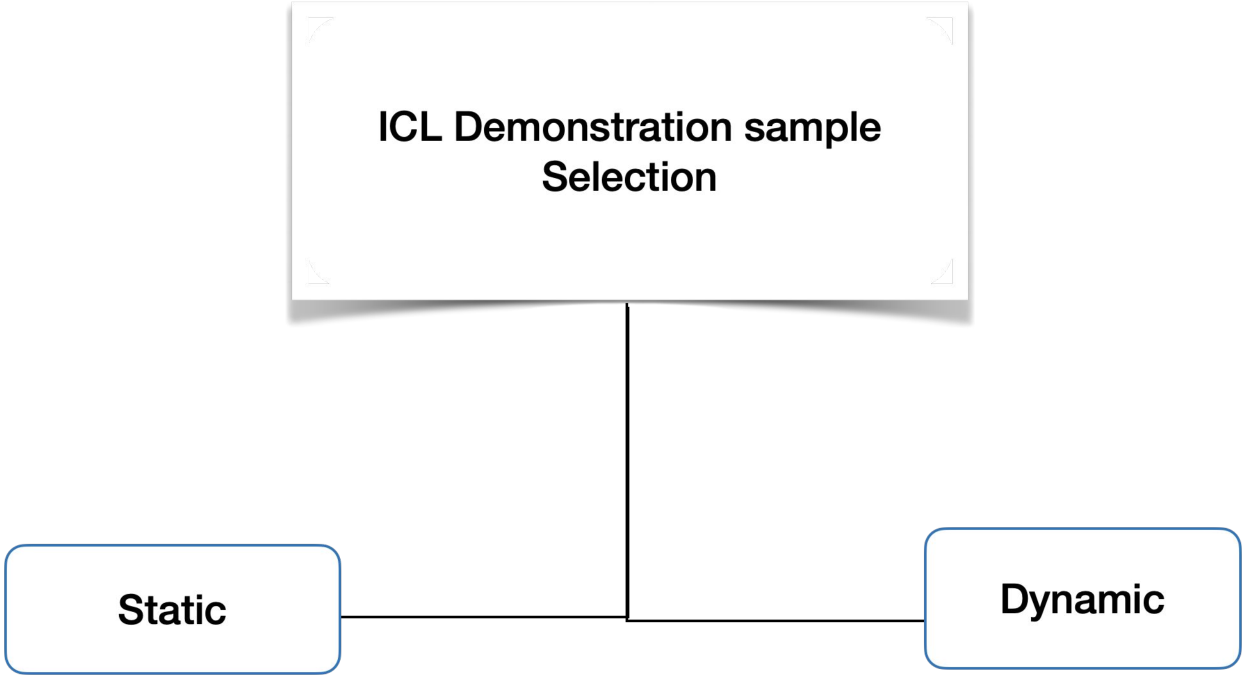
Fig: Block Diagram of ICL

Exemplars / In-context examples / demonstration samples

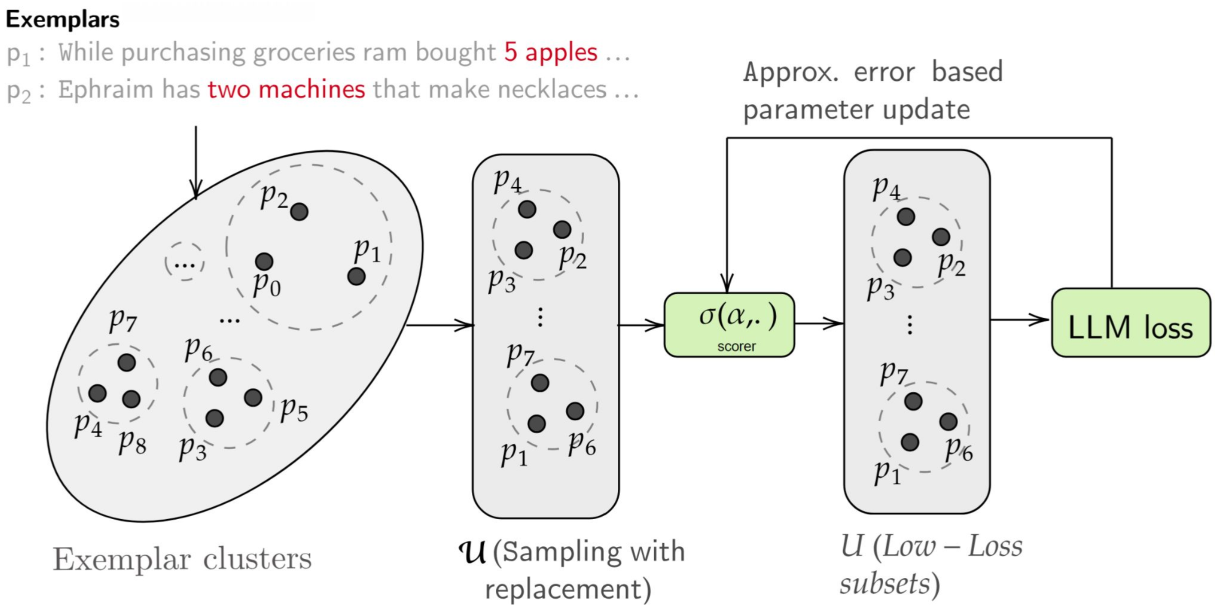
<Question, Explanation, Answer>



ICL types



Explore-Exploit Paradigm



* accepted at **EMNLP-main (long) 2024** (*EXPLORA: Efficient Exemplar Subset Selection for Complex Reasoning*)

Loss Modeling

Subset of k Exemplars ($S \subseteq \mathcal{S}$)

$$\overbrace{\sigma(\vec{\alpha}, S) = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} \alpha_i (x_i \in S) E_{ij}}^{\text{Loss modelling function; Approximating } L(S,\mathcal{V})} \qquad (1)$$

ith exemplars contribution, low if important exemplar

Any transformer based encoder

$$E_{ij} = \frac{\phi(x_i)^T \phi(u_j)}{\|\phi(x_i)\| \|\phi(u_j)\|}$$

ith exemplar, $x_i \in S$

ith validation sample, $u_i \in \mathcal{V}$

Efficient Estimation of parameters

Estimating loss here involves LLM calls and equivalent to arm pulling

Update parameters to reduce approximation error

set of l subsets at timestep t with lowest validation loss

Validation Set

$$\mathcal{L}(\vec{\alpha}; U_t, V_t) = \sum_{S \in U_t} (L(S, \mathcal{V}) - \sigma(\vec{\alpha}, S))^2 + \sum_{S' \sim V_t} (L(S', \mathcal{V}) - \sigma(\vec{\alpha}, S'))^2$$

Remaining subsets; $V_t \leftarrow \mathcal{U} \setminus U_t$, where $\mathcal{U} \subset \mathcal{S}$

Negative samples from high-loss set V_t

Results and Analysis

Method	GSM8K	AquaRat	TabMWP	FinQA	StrategyQA
		GPT-3.5-turbo			
dynamic					
KNN (Rubin et al., 2022)	53.45	51.96	77.07	51.52	81.83
KNN (S-BERT) (Rubin et al., 2022)	53.07	52.75	77.95	52.65	81.83
MMR (Ye et al., 2023b)	54.36	51.18	77.32	49.87	82.86
KNN+SC (Wang et al., 2023c)	80.21	62.59	83.08	54.49	83.88
MMR+SC (Wang et al., 2023c)	78.01	59.45	81.36	50.74	83.88
PromptPG (Lu et al., 2023b)	-	-	68.23	53.56	-
static					
Zero-Shot COT (Kojima et al., 2023)	67.02	49.60	57.10	47.51	59.75
Manual Few-Shot COT (Wei et al., 2023)	73.46	44.88	71.22	52.22	73.06
Random	67.79	49.80	55.89	53.70	81.02
PS+ (Wang et al., 2023b)	59.30	46.00	-	-	-
Auto-COT (Zhang et al., 2023b)	57.10	41.70	-	-	71.20
GraphCut (Iyer and Bilmes, 2013)	66.19	47.24	60.45	52.31	80.00
FacilityLocation (Iyer and Bilmes, 2013)	68.61	48.43	67.66	36.79	81.63
LENS (Li and Qiu, 2023)	69.37	48.82	77.27	54.75	79.79
LENS+SC (Li and Qiu, 2023)	79.37	57.87	80.68	60.06	82.24
Our Approach					
EXPLORA	77.86(▲12.24%) †	53.54(▲9.67%)†	83.07(▲7.51%) †	59.46(▲8.60%) †	85.71(▲5.63%) †
EXPLORA+SC	**86.35**(▲24.48%) ‡	63.39(▲29.84%) ‡	85.52(▲10.68%) ‡	64.52(▲17.84%) ‡	87.14 (▲9.21%)†
EXPLORA+KNN+SC	85.14 (▲22.73%)‡	62.20(▲27.41%)‡	86.29(▲12.39%) ‡	**65.12**(▲18.94%) ‡	**88.37**(▲10.75%)†
EXPLORA+MMR+SC	86.13(▲24.16%) ‡	**63.78**(▲30.64%) ‡	**86.96**(▲12.54%)‡	64.60(▲17.99%) ‡	87.55(▲9.73%)†
		GPT-4o			
LENS (Li and Qiu, 2023)	76.19	64.56	86.34	69.31	92.85
EXPLORA	93.63	69.29	90.12	72.71	95.10

Table: Results across datasets in transfer setting using gpt-3.5-turbo with exemplars selected from Mistral-7b.

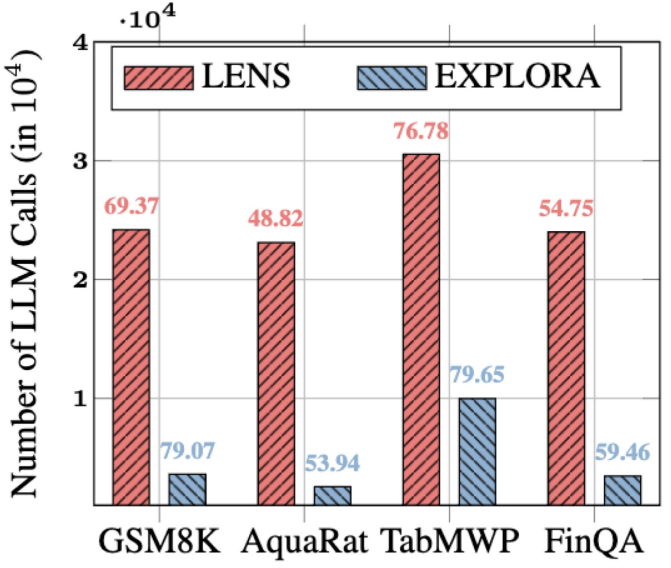
Prompt Transfer Works Well

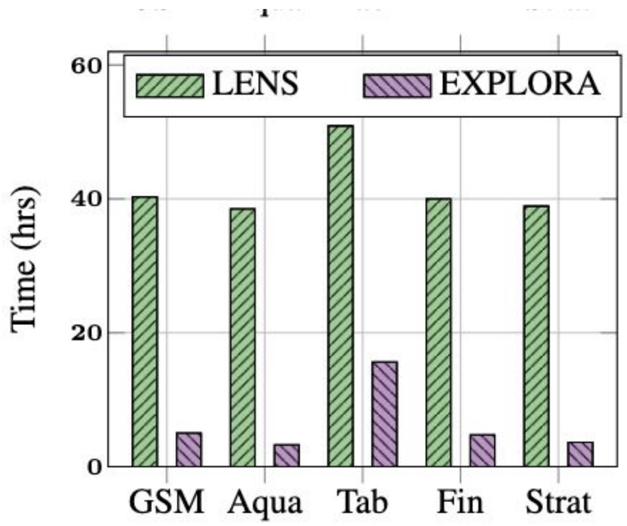
Method	T	GSM	Aqua	Tab	Fin
EXP	L	79.07	53.94	79.65	54.66
	M	77.86	53.54	77.41	59.46
EXP+SC	L	85.82	63.78	86.76	61.16
	M	86.35	63.39	85.52	64.52
EXP+KNN+SC	L	85.89	64.17	85.74	63.64
	M	85.14	62.20	86.29	65.12
EXP+MMR+SC	L	86.20	62.99	87.81	64.60
	M	86.13	63.78	86.96	64.60

EXPLORA is more Robust

Datasets	GSM	Aqua	Tab	Fin
Zero-Shot COT	±5.18	±7.08	±1.84	±4.50
Few-Shot COT	±4.48	±12.03	±1.66	±4.76
KNN	±3.76	±5.49	±1.27	±4.17
MMR	±4.00	±10.53	±1.68	±6.10
Graph Cut	±6.38	±8.18	±2.03	±5.29
Facility Location	±4.23	±6.71	±1.74	±4.94
LENS	±5.04	±6.67	±1.72	±5.81
EXPLORA	±**3.39**	±**4.93**	±1.45	±**3.41**

EXPLORA is resource Efficient





THANK YOU
FOR
YOUR ATTENTION!!!

https://github.com/kiranpurohit/

@kiranpurohit08