# Enhancing User Behavior Alignment by Input-Level Model Cooperation and Model-Level Parameter Optimization

Team members:    Yingyi Zhang
                 Zhipeng Li
                 Zhewei Zhi
Supervisor:   Xianneng Li

# Content

01 | Introduction

# Track 3: User Behavior Alignment



User behavior

- Click
- Query
- Buy
- ⋮
- Review

Hidden regularity

- Also buy
- Satisfaction
- Preference
- ⋮
- Intention

Input
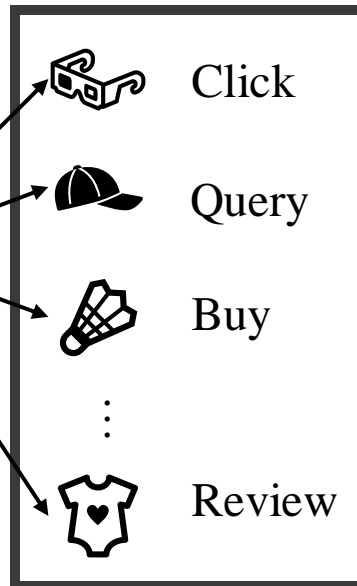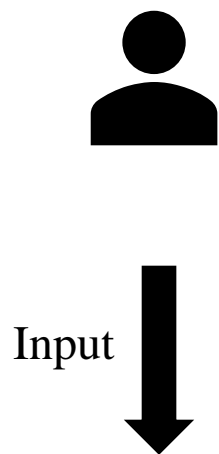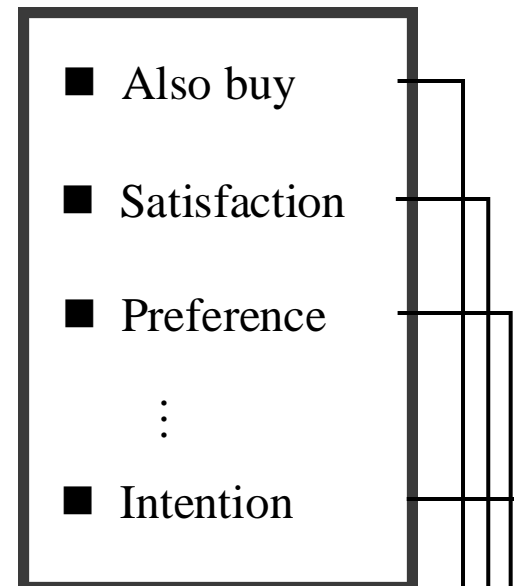
Alignment

Business insight

- Substitute or complementary
- Sentiment of the review
- Retrieval related products
- ⋮
- Ranking products
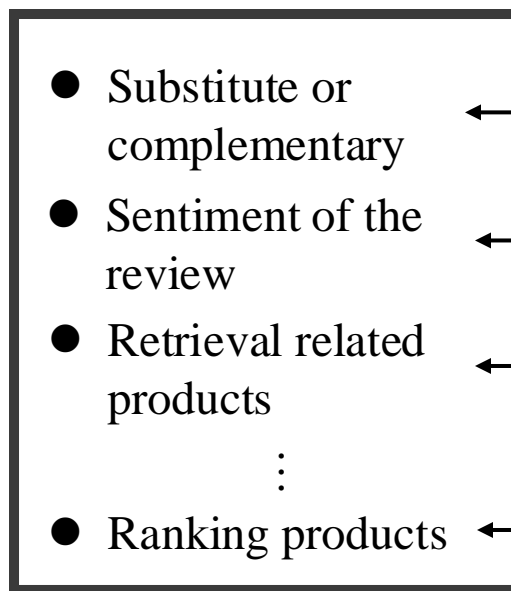
# Examples

## Multiple choice

Which of the following statements best describes the relation from query "waterpik" to query "long sleeve crop tops for women"?
0. irrelevant
1. substitute
2. complement
3. narrowing
Answer: 0

Instructions: Evaluate the following product review on a scale of 1 to 5, with 1 being very negative and 5 being very positive.
Input: a review text for a(n) cereal product
I love this Meusli. Its very simple. Im sure I could mix up my own, but its easier to just scoop it out for use. I haven't tried cooking it. I usually put it in a bowl with a little almond milk and let it sit for about 10 minutes, which I think is what the package directions say. Id like to try cooking it sometime. Its very good. Its currently my favorite cereal and, I think, much better for you than other packaged breakfast cereals.
Output:
Answer: 5

## Ranking

A user has made a query with keyword 'jeep liberty lift'. Given the following numbered list of 5 products, please rank the products according their relevance with the query.
Product List:
1. Supreme Suspensions - Front Leveling Kit for 2002-2007 Jeep Liberty KJ and 2008-2012 Jeep Liberty KK 2.5" Front Lift High-Strength Carbon Steel Strut Spacers 2WD 4WD
2. Rough Country 2.5" Lift Kit for 2007-2018 Jeep Wrangler JK 4DR - 67930
3. Rough Country 2.5" Lift Kit (fits) 1997-2006 Jeep Wrangler TJ LJ | 6 CYL | N3 Shocks | Suspension System | 653.20
4. Supreme Suspensions - Full Lift Kit for 2008-2012 Jeep Liberty KK 2.5" Front Strut Spacers + 2" Rear Spring Spacers High-Strength Carbon Steel Lift Kit 2WD 4WD PRO KIT
5. TeraFlex 1251000 2.5" Lift Kit (JK 4 Door with All (4) 2.5" Shock)
You should output a permutation of 1 to 5. There should be a comma separating two numbers. Each product and its number should appear only once in the output. Only respond with the ranking results. Do not say any word or explanations.
Output: [4,1,3]

# Examples

A user on an online shopping website has just purchased a product 'Flanged End Cap for L-Track'. The following numbered list contains 15 products. Please select 3 products from the list that the user may also purchase.

Product List:

1. Blacktop &amp; Roof Patch,10.1 Oz (Pack of 6)
2. DEWALT D25553K 1-9/16-Inch Spline Combination Hammer Kit
3. XtremepowerUS 10pc 1&quot; Dr.Deep Impact Cr-V Socket Set - MM (Black)
4. Forney 60224 Mini Rotary File Cutter Set, 1/8-Inch Shaft, 3-Piece
5. Voltec 08-00616 1400-Watt Halogen Pro Worklight, 7-Foot, Blue &amp; Yellow
6. 3.5&quot; Acrylic Lens Rimless 2x Magnifying Glass w/2 LEDs - Great for Basic Inspections, Perfect for Crafts &amp; Hobbies!
7. Icicle Solar Christmas String Lights, 15.7ft 8 Light Modes 20 LED Water Drop Fairy String Lighting for Outdoor &amp; Indoor, Home, Patio, Lawn, Garden, Party, and Holiday Decorations (Warm White)
8. Blue Sea Systems 187 Series, 285 Series &amp; Klixon Circuit Breakers
9. Ridgid 59832 Die Head Post
10. Johnson Level &amp; Tool 175-L Post Level
11. Wire Loom Black 20' Feet 1&quot; Split Tubing Hose Cover Auto Home Marine by Nippon America
12. Brinks 7462-619 Hampton 3-Light Camille Bath Vanity Light, Satin Nickel
13. Platform Stepladder, 7 ft. 9in, 330 lb.
14. Self-Adhesive Stress Crack Tape Textured Roll
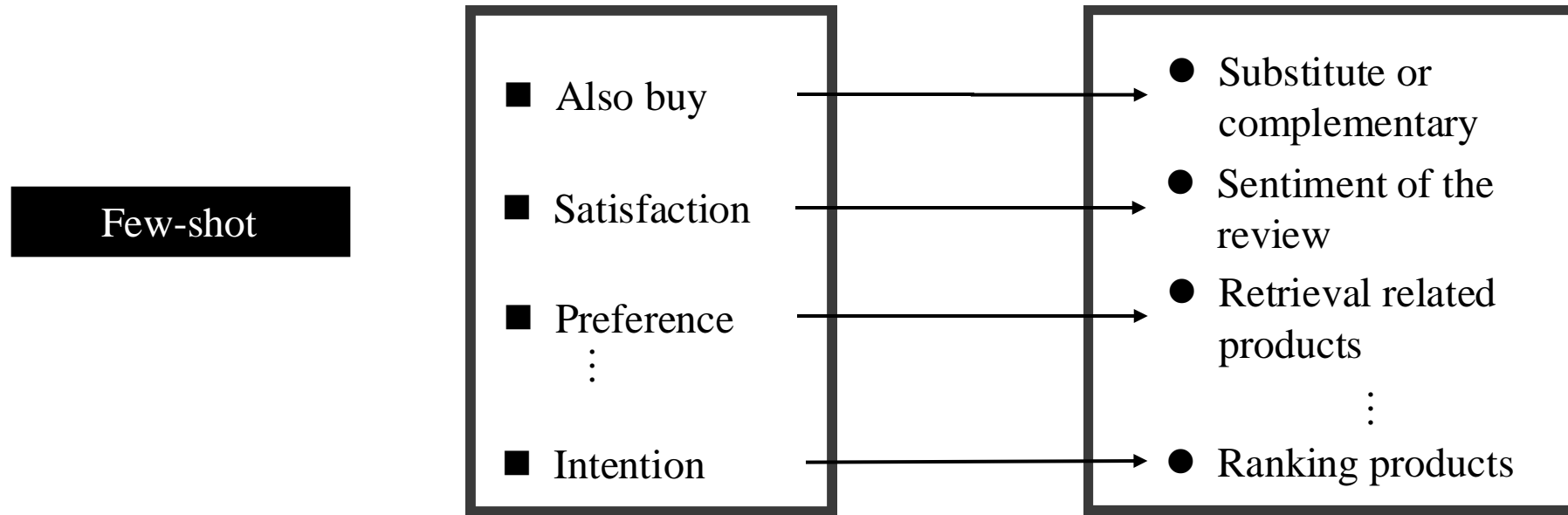15. SHURFLO (255-313) 1/2&quot; Twist-On Pipe Strainer

You should output 3 numbers that correspond to the selected products. There should be a comma separating every two numbers. Only respond with the results. Do not say any word or explanations.

Output: [8, 11, 15]

# 02 Challenge and Motivation

# Cahllenge

| Task | Number of task | Number of samples |
|---|---|---|
| Few-shot | 5 | 27 |
| Zero-shot | 10 | 3949 (1576 in round 1 and 2373 in round 2) |

**Few-shot**

- Also buy → ● Substitute or complementary
- Satisfaction → ● Sentiment of the review
- Preference → ● Retrieval related products
- Intention → ● Ranking products

**Zero-shot**

**Challenge:** nobody knows the new task of user behavior

# Motivation

| Task | Number of task | Number of samples |
|------|---------------|-------------------|
| Few-shot | 5 | 27 |
| Zero-shot | 10 | 3949 (1576 in round 1 and 2373 in round 2) |
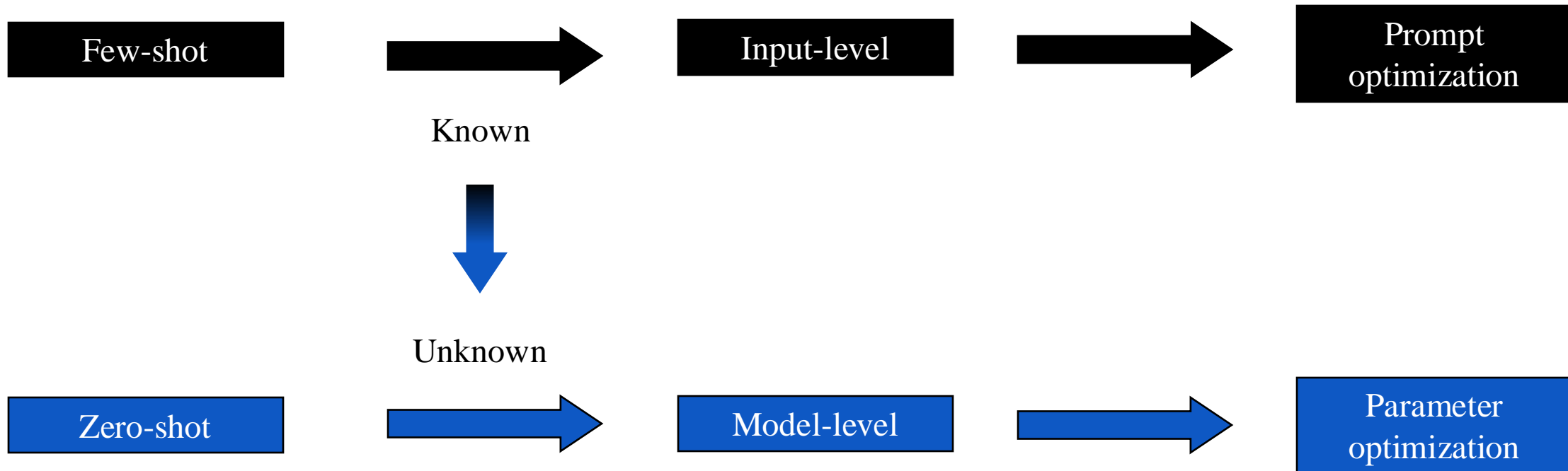
# Motivation

| Task | Number of task | Number of samples |
|---|---|---|
| Few-shot | 5 | 27 |
| Zero-shot | 10 | 3949 (1576 in round 1 and 2373 in round 2) |
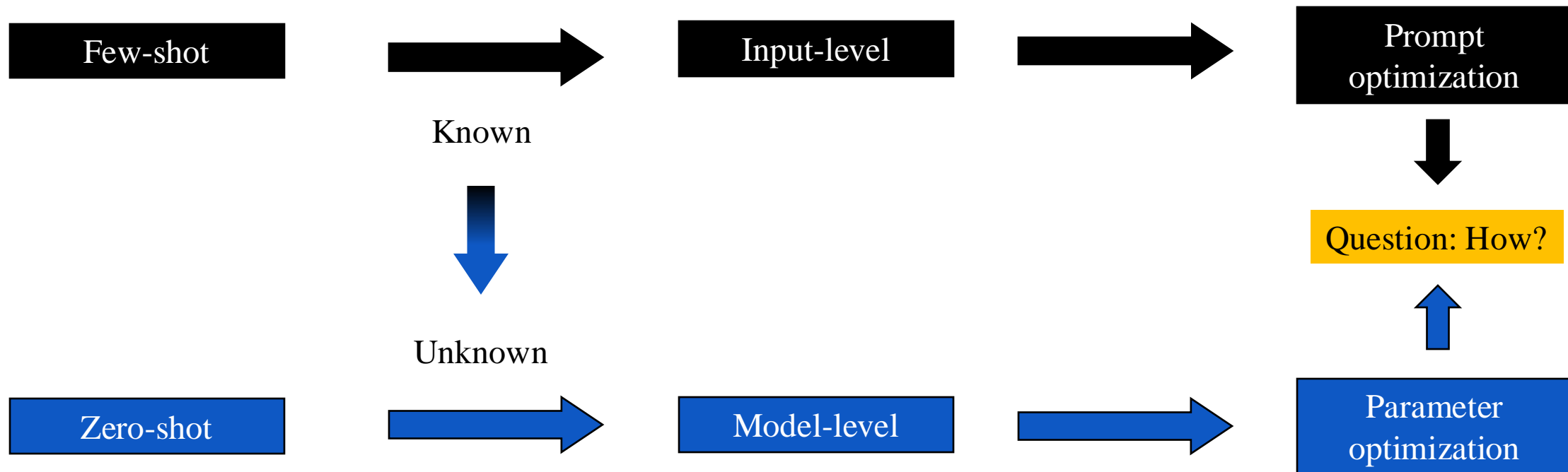
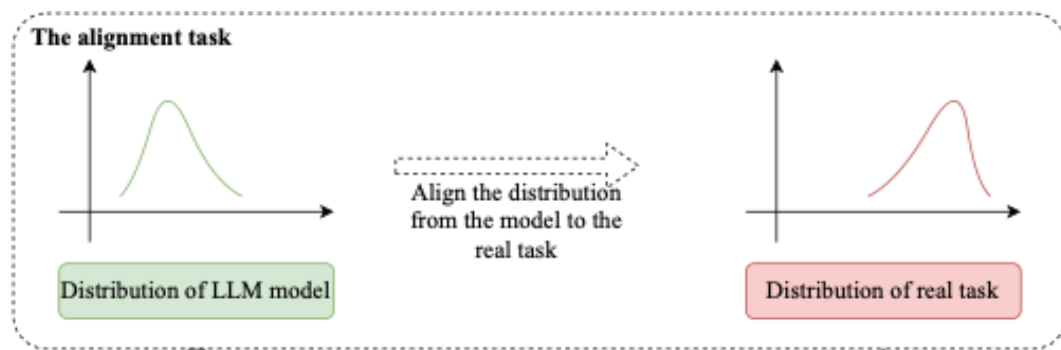03 | Overview of Our Solution
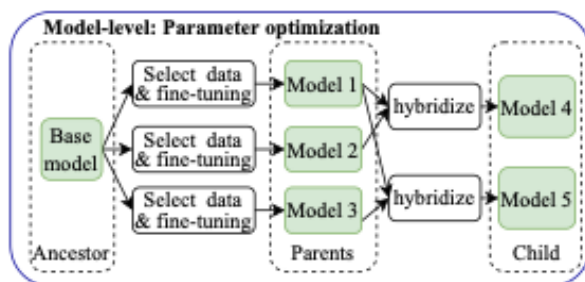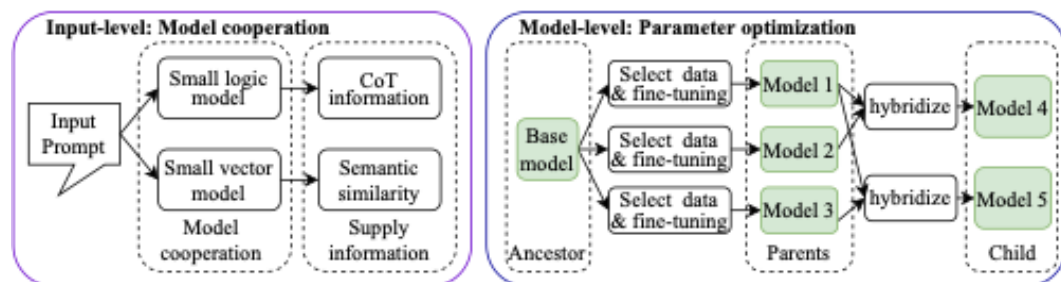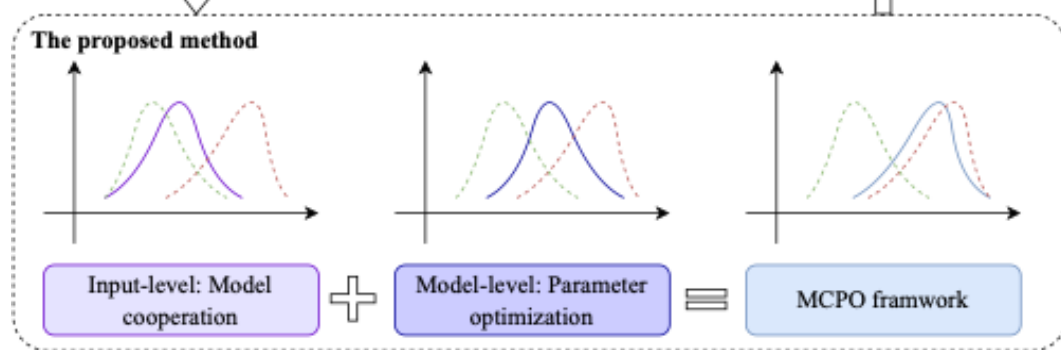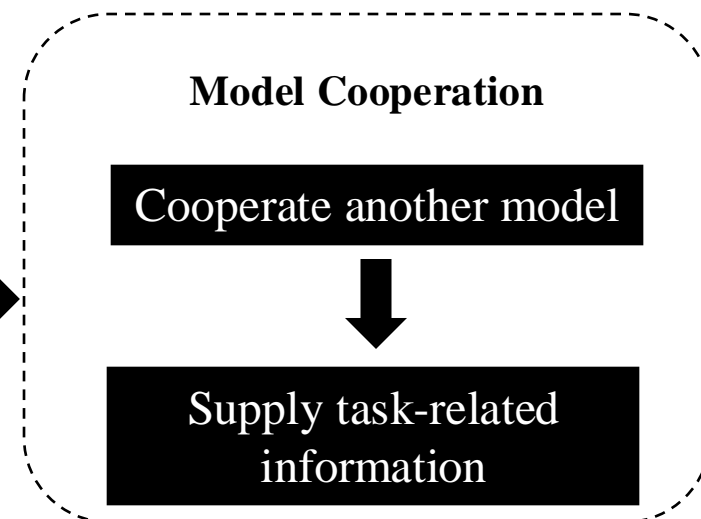
# Overall solution

04

04 | Spotlight

# Model selection
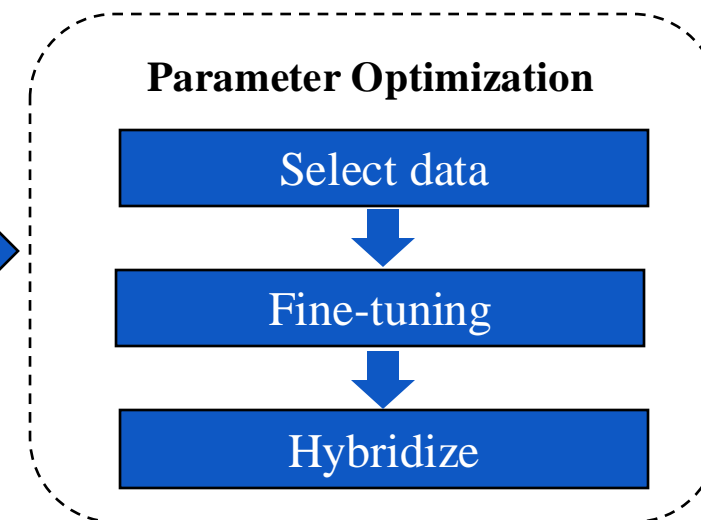
The difference between the effects of different models is very large, in general:

➤ The larger the model parameters, the better the model effect

➤ Models with a similar number of parameters will be much better when fine-tuned

**Table 3: Results of model selection**

| Model | Score | Multiple Choice Score | Retrieval Score | Generation Score | Ranking Score |
|---|---|---|---|---|---|
| ecellm-L | 0.5310 | 0.5543 | 0.4230 | 0.4953 | 0.7761 |
| ecellm-M | 0.5185 | 0.4970 | 0.4556 | 0.5568 | 0.7632 |
| llama3-8b | 0.4198 | 0.4549 | 0.3283 | 0.3269 | 0.6922 |
| Wizard-Vicuna-30B-Uncensored-AWQ | 0.3605 | 0.2984 | 0.3010 | 0.4844 | 0.6641 |
| Qwen1.5-14B | 0.2069 | 0.2003 | 0.2792 | 0.0267 | 0.5825 |
| llama3-70B-AWQ | **0.6657** | **0.6406** | **0.7667** | **0.5755** | **0.8333** |

# Prompt engineering

| Task | Prompt | Score |
|------|--------|-------|
| MC | You are a helpful online shopping assistant. Please answer the following question about online shopping and follow the given instructions. n\n | 0.697 |
| | You are a helpful online shopping assistant. Please answer the following question about online shopping and follow the given instructions. Please think cautiously. The answers or reponses are very important.n\n | 0.704 |
| | complement → 'Users who buy products 1 may also buy products 2'<br>irrelevant → products 1 + ' and ' + products2 + ' are irrelevant'<br>substitute → products1 + ' and ' + products2 + ' are similar'<br>narrowing → products1 + ' is a further search of ' + products2 | 0.710 |

# Prompt engineering

| Task | Prompt | Score |
|------|--------|-------|
| Retrieval | You are a helpful online shopping assistant. Please answer the following question about online shopping and follow the given instructions. n\n | 0.811 |
| | This is a retrieval question. You are a highly skilled online shopping assistant and a professional product retrieval expert. Your goal is to help consumers quickly and accurately identify products that meet their specific needs. You provide a clear and concise list of retrieval results, including the product name, key attributes, and how they meet the requirements. Please analyze the following request and deliver accurate retrieval results. | 0.834 |
| | Based on your previous knowledge then generate the answer. This is a retrieval question. You are a highly skilled online shopping assistant and a professional product retrieval expert. Your goal is to help consumers quickly and accurately identify products that meet their specific needs. You provide a clear and concise list of retrieval results, including the product name, key attributes, and how they meet the requirements. Please analyze the following request and deliver accurate retrieval results. | 0.838 |
| | Based on your previous knowledge then generate the answer as clearly as possible. This is a retrieval question. You are a highly skilled online shopping assistant and a professional product retrieval expert. Your goal is to help consumers quickly and accurately identify products that meet their specific needs. Plese provide a clear and concise list of retrieval results. Please analyze the following request and deliver accurate retrieval results.\n\n | 0.840 |

# Prompt engineering

| Task | Condition | Prompt | Score |
|---|---|---|---|
| Generation | - | You are a helpful online shopping assistant. Please answer the following question about online shopping and follow the given instructions. \n\n | 0.496 |
| | product description | You are a marketing specialist. Write a detailed and persuasive product description based on the following instruction.\n\n | |
| | customer Review | You are a satisfied customer. Write a genuine and detailed review. Mention the features you liked the most and how the product met your expectations.\n\n | |
| | recommendation | You are a personal shopping assistant. Generate personalized product recommendations for a customer based on their recent browsing history and purchase patterns based on the following instruction.\n\n | 0.608 |
| | email | Social Media Post:\nGenerate a social media post for an e-commerce product. Include an engaging opening statement, key features, relevant hashtags, and a call to action.\n\n | |
| | AQ | You are a customer service representative. Generate a answer of the question based on the following instruct. Ensure the answers are clear and helpful to potential customers.\n\n | |
| | else: | You are assistant in online e-commerce. Now you need to answer the generation task based on the following instruct.\n\n | |
| | | Based on your previous knowledge then generate the answer.<br>Response (limit to 30 words) | 0.630 |

# RAG Method For Ranking Task

```
You are an intelligent shopping assistant that can rank products based on their relevance to the query. The following numbered list contains 5 prod
ucts. Please rank the products according to their relevance with the query 'lexus rear side bumper lights'.
Product List:
1. Marsauto 194 LED Light Bulb 6000K 168 T10 2825 5SMD LED Replacement Bulbs for Car Dome Map Door Courtesy License Plate Lights (Pack of 10)
2. LivTee Truck Tailgate Light Bar 60" LED Strip with Red Running Brake White Reverse Red Turning Signals Lights - IP68 Waterproof
3. DSparts Rear Left Side Marker Bumper Light Fits FOR 2004-2009 Lexus RX330 RX350 RX400H
4. Nilight 2PCS 18W 1260lm Spot Driving Fog Light Off Road Led Lights Bar Mounting Bracket for SUV Boat 4" Jeep Lamp,2 years Warranty
5. Motor Trend 923-GR Gray FlexTough Contour Liners-Deep Dish Heavy Duty Rubber Floor Mats for Car SUV Truck & Van-All Weather Protection, Universa
l Trim to Fit


[External information]
Here are some products if user input the query "lexus rear side bumper lights":
1.DSparts Rear Left Side Marker Bumper Light Fits FOR 2004-2009 Lexus RX330 RX350 RX400H;
2.GTINTHEBOX 3D Optic Red LED Rear Bumper Reflectors Brake Tail Lights Lamps for Lexus RC250 RC350 is-F GX470 Toyota REIZ Sienna Matrix Venza Avalo
n 2Pcs;
3.Reflector Lens & Housing Rear Passenger Side Right for 06-13 Lexus IS250 IS350;Rareelectrical NEW LEFT REAR SIDE MARKER LIGHT COMPATIBLE WITH LEX
US RX350 2007-2009 LX2861102 819100E010 81920-0E010 819200E010 LX2860102;TYC Left Side Marker Light Assembly Compatible with 2004-2009 Lexus RX 330
/350;TYC Right Side Marker Light Assembly Compatible with 2004-2009 Lexus RX 330/350;REVi MotorWerks DEPO Smoke 4 Pieces Front + Rear Bumper Side M
arker Lights Combo Compatible with 2000-2005 Lexus IS300 / IS200
The following items are very important and have a strong relevance to the query, and ranked by relevance.
3. DSparts Rear Left Side Marker Bumper Light Fits FOR 2004-2009 Lexus RX330 RX350 RX400H
2. LivTee Truck Tailgate Light Bar 60" LED Strip with Red Running Brake White Reverse Red Turning Signals Lights - IP68 Waterproof
1. Marsauto 194 LED Light Bulb 6000K 168 T10 2825 5SMD LED Replacement Bulbs for Car Dome Map Door Courtesy License Plate Lights (Pack of 10)
4. Nilight 2PCS 18W 1260lm Spot Driving Fog Light Off Road Led Lights Bar Mounting Bracket for SUV Boat 4" Jeep Lamp,2 years Warranty
5. Motor Trend 923-GR Gray FlexTough Contour Liners-Deep Dish Heavy Duty Rubber Floor Mats for Car SUV Truck & Van-All Weather Protection, Universa
l Trim to Fit
You should output a permutation of 1 to 5 reference to the order of [External information]. There should be a comma separating two numbers. Each pr
oduct and its number should appear only once in the output. Only respond with the ranking results. Do not say any word or explanations.
Output: 3,1,2,4,5
```

**External Informaion**

Idea: Find the relevant results of the query term and use it to search the candidate products

Steps:
1. Find relevant results according to the query term
2. Vectorized related results with individual options, problems with options do inner products, sort recalls

# RAG Method For Ranking Task

## Table 2: Result of applying semantic similarity

| Ranking RAG | RKQ Score |
|---|---|
| baseline (llama3-70B) | 0.737 |
| +system_prompt | 0.839 |
| $k = 1$ | 0.890 |
| $k = 3$ | **0.901** |
| $k = 5$ | 0.888 |

| Task | Prompt | Score |
|---|---|---|
| Ranking | You are a helpful online shopping assistant. Please answer the following question about online shopping and follow the given instructions. n\n | 0.888 |
| | Based on your previous knowledge then generate the answer. + prompt | 0.901 |
| | Prompt + In this task, each question is associated with a requirement and a list of candidate items, and the model is required to re-rank all items according to how each item satisfies the requirement.\n\n | 0.914 |
| | Prompt + Only respond with the ranking results. Do not say any word or explanations | 0.918 |

# Model Fine-tuning

## More fine-tuning data is not always better, but appropriate!!!

| Model | Score | Multiple Choice Score | Retrieval Score | Generation Score | Ranking Score |
|---|---|---|---|---|---|
| baseline | 0.665 | 0.640 | 0.766 | 0.575 | 0.833 |
| $N = 2000$ | 0.693 | 0.687 | 0.789 | 0.543 | **0.901** |
| $N = 4750$ | 0.710 | 0.692 | **0.826** | 0.584 | 0.883 |
| $N = 6000$ | **0.715** | 0.697 | 0.811 | **0.608** | 0.888 |
| $N = 6100$ | 0.699 | 0.680 | 0.795 | 0.586 | 0.897 |
| $N = 10000$ | 0.696 | 0.676 | 0.804 | 0.586 | 0.863 |
| $N = 20000$ | 0.695 | 0.700 | 0.768 | 0.551 | 0.873 |
| Hybridized model (6000-20000) | 0.709 | **0.703** | 0.796 | 0.584 | 0.868 |

# Model hybridize
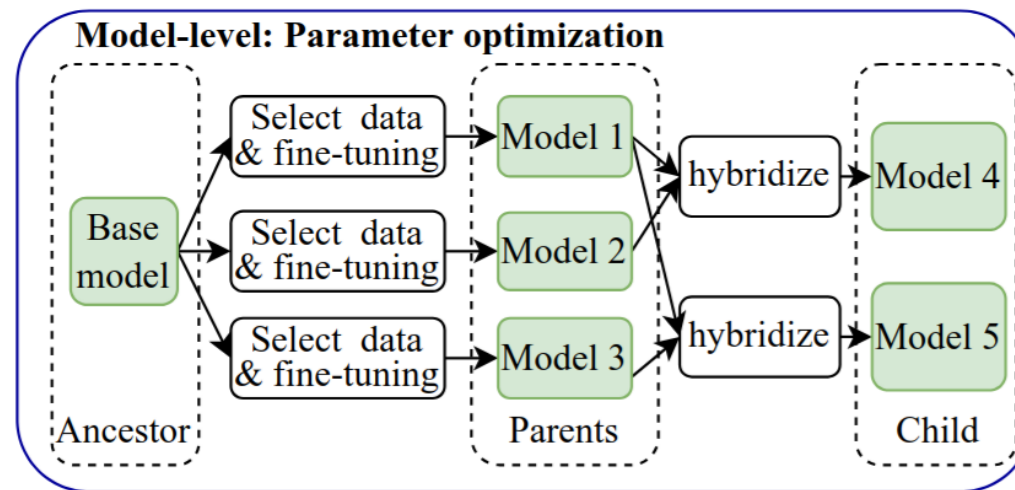


| Generation | Model | Score | MCQ Score | RTQ Score | GENQ Score | RKQ Score |
|---|---|---|---|---|---|---|
| Ancestor | Phi-3b($base$) | 0.4152 | 0.3306 | 0.2750 | 0.4872 | 0.7372 |
| Parent | Phi-3b($N = 2000$) | 0.4680 | 0.5096 | 0.0440 | 0.3424 | 0.8929 |
|  | Phi-3b($N = 10000$) | 0.4597 | 0.4710 | 0.0980 | 0.4026 | 0.8448 |
|  | Phi-3b($N = 20000$) | 0.4631 | 0.5083 | 0.0190 | 0.3164 | 0.9184 |
| Child | Phi-3b($N = 2000 + N = 20000$) | 0.4658 | 0.5080 | 0.0310 | 0.3347 | 0.9056 |
|  | Phi-3b($N = 20000 + N = 10000$) | **0.4699** | 0.4890 | 0.0730 | 0.4048 | 0.8745 |
| Ancestor | Qwen2-7b($base$) | 0.5187 | 0.5323 | 0.3530 | 0.4864 | 0.6760 |
| Parent | Qwen-7b($N = 20000$) | 0.6551 | 0.6826 | 0.5060 | 0.4603 | 0.9165 |
|  | Qwen-7b($N = 101275$) | 0.6441 | 0.6743 | 0.5150 | 0.4636 | 0.8630 |
| Child | Qwen-7b($base + N = 20000$) | 0.6085 | 0.6540 | 0.4420 | 0.4795 | 0.7676 |
|  | Qwen-7b($N = 20000 + N = 101275$) | **0.6719** | 0.7153 | 0.5730 | 0.4631 | 0.8495 |
| Ancestor | LLaMA3-70b($base$) | 0.6071 | 0.5543 | 0.5920 | 0.5115 | 0.8762 |
| Parent | LLaMA3-70b($N = 20000$) | 0.7393 | 0.7530 | 0.7420 | 0.4874 | 0.9472 |
|  | LLaMA3-70b($N = 101275$) | 0.7498 | 0.7716 | 0.7690 | 0.4852 | 0.9299 |
| Child | LLaMA3-70b($base + N = 101275$) | 0.6213 | 0.7543 | 0.6540 | 0.4999 | 0.3107 |
|  | LLaMA3-70b($N = 20000 + N = 101275$) | **0.7518** | 0.7733 | 0.7800 | 0.4852 | 0.9459 |

Evaluation on offline dataset

05 | Conclusion

# Conclusion

**Input-Level**

**Prompt optimization**

**Model Cooperation**

Cooperate another model

↓

Supply task-related information
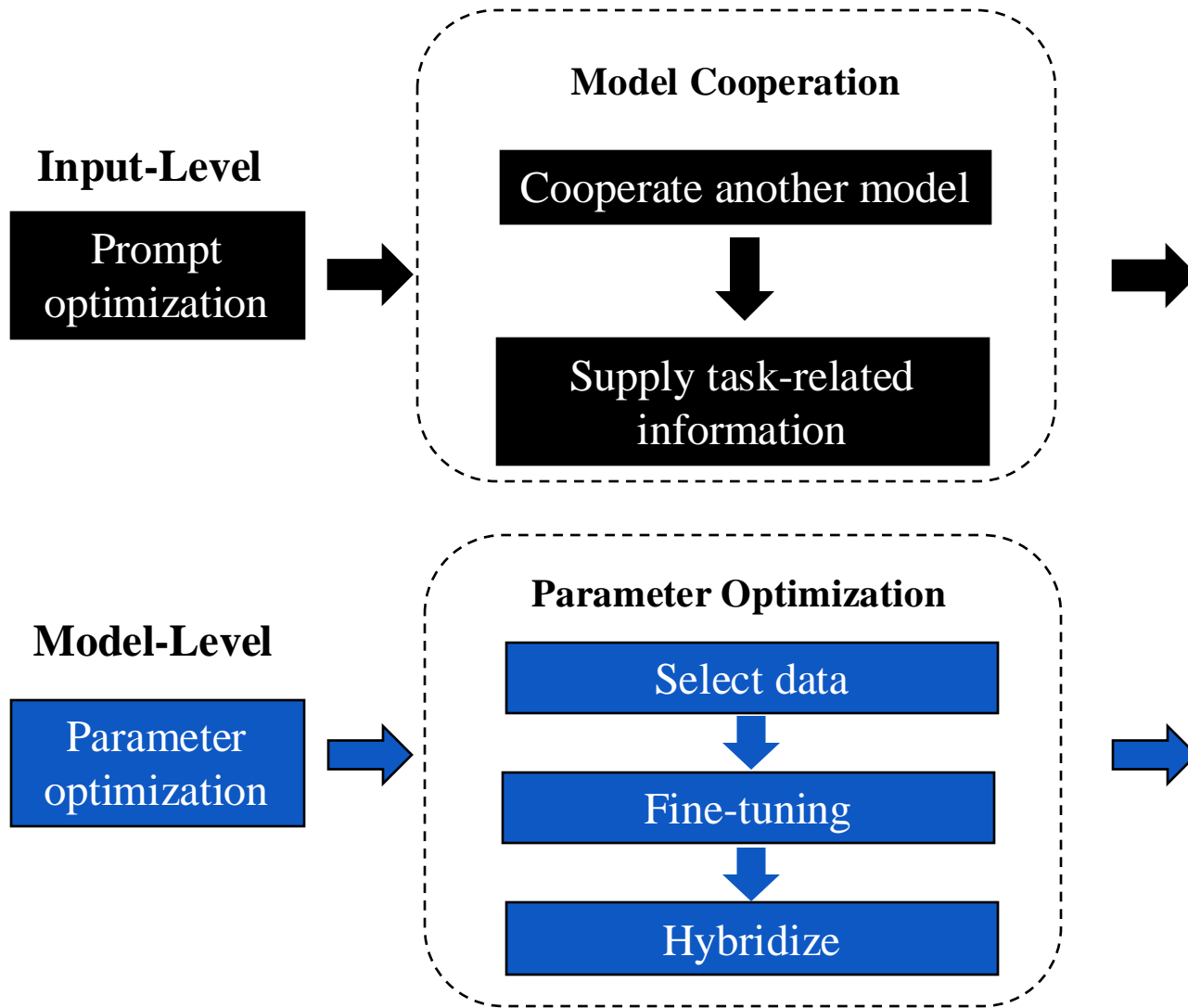
1. **Not one model to serve all:** Cooperate different model can enhance LLM performance in specific task.
2. **Explicit present knowledge is important:** Our CoT and semantic similarity shown that explicit present knowledge and input it into LLM is very effective.

**Model-Level**

**Parameter optimization**

**Parameter Optimization**

Select data

↓

Fine-tuning

↓

Hybridize

1. **Not the more data the better:** In fine-tuning, we find that 6000 version is better than other.
2. **Hybridize can emerge new ability:** The 6000+2000 version can perform a better MCQ ability.

# Thanks

Team members:   Yingyi Zhang    yingyizhang@mail.dlut.edu.cn

Zhipeng Li    lizhipeng@mail.dlut.edu.cn

Zhewei Zhi    dzwer@mail.dlut.edu.cn

Supervisor:   Xianneng Li    xianneng@dlut.edu.cn

大连理工大学
DALIAN UNIVERSITY OF TECHNOLOGY