多角色机制Prompt设计

|  |
| --- |
| **## Role**  XXXX. Analyze the conversation context and the roles of participants to determine the correct classification. |
| **### Participants**   1. **Issue Raiser**   - Describes specific issues or bugs with details such as environment and error logs.  - Requests features or suggests improvements.  - Expresses concerns or opinions about existing functionality.  - Provides steps to reproduce issues and relevant logs.   1. **Commenter**   - Shares similar experiences or confirms issue prevalence.  - Adds supplementary information or explains issue context.  - Suggests potential solutions or theoretical insights.  - Engages in technical discussions or responds to comments.   1. **Member**   - Assesses technical feasibility and validates issue existence.  - Assigns tasks based on issue importance or urgency.  - Guides discussions and clarifies issue focus.  - Adds labels, assigns personnel, and updates issue status.   1. **Collaborator**   - Proposes technical solutions, including code or system analyses.  - Submits pull requests (PRs) to address issues directly.  - Analyzes and validates proposed fixes.  - Offers forward-looking advice to enhance project quality.   1. **Contributor**   - Provides historical context or connections to similar issues.  - Reproduces issues and confirms conditions for replication.  - Verifies fixes to ensure no new problems are introduced.  - Offers supplementary insights or suggestions.   1. **Owner**   - Makes final decisions within technical and resource constraints.  - Prioritizes issues and plans tasks for short- and long-term goals.  - Proposes long-term solutions or future-oriented designs.  - Reviews and approves fixes submitted by others.  - Closes resolved issues or converts them into actionable tasks. |

混合自我反思机制Prompt设计

|  |
| --- |
| 1. **Read the Conversation and Repository Description:**   Begin by reviewing the repository description thoroughly to understand the project’s purpose, supported features, and scope. Keep this context in mind as you examine each round of the conversation. If the problem discussed clearly falls outside the repository’s described capabilities or instructions, consider “other” or “question.”   1. **For Each Round:**  * Classify the round into one of the five categories. * self\_reflection: Justify your classification’s accuracy. Confirm whether it matches the repository’s description and consider alternatives. If unsure, reflect on the possibility of user misunderstanding vs. repository code issues.  1. **Final Self-Reflection Across Rounds:**  * After classifying all rounds, review the entire conversation. * Check for consistency:   + Does the chosen final category align with the majority of evidence?   + Are there any contradictions between rounds?   + Could the issue be a misunderstanding (question) rather than a repository problem (error/performance/deployment)?   + If the issue is irrelevant to the repository description, finalize as \*\*other\*\*.  1. **Finalize the Classification** |

Zero-Shot方法的输入与输出

|  |  |
| --- | --- |
| Input | Output |
| * "The repository fails to install correctly. When I run pip install, it throws a dependency error about PyTorch." * "Have you checked if your Python version matches the requirements in the README? The repository requires Python 3.8+." * "I'm using Python 3.7. Let me update and retry."  "After upgrading Python, it works now. Thanks!" | **Classification: deployment**  Evidence: "The installation failed due to a dependency error, later resolved by aligning with the repository requirements." |

Issue 对话分类框架的输入与输出

|  |  |
| --- | --- |
| Input | Output |
| Round 1: Issue Raiser: "The repository fails to install correctly. When I run pip install, it throws a dependency error about PyTorch." | Role Analysis: Issue Raiser describes the problem without confirming adherence to prerequisites.  **Classification: question.**  Evidence: "Dependency error points to a likely user misconfiguration rather than a defect in the repository's deployment process."  Self-Reflection: "This might not be a deployment issue; user could be using an unsupported Python version. More details needed." |
| Round 2: Member: "Have you checked if your Python version matches the requirements in the README? The repository requires Python 3.8+." Issue Raiser: "I'm using Python 3.7. Let me update and retry." | Role Analysis: Member identifies the likely root cause and provides guidance based on repository requirements.  Classification: question.  Evidence: "Member directs the Issue Raiser to check their Python version. The user admitted using an unsupported Python version (3.7)."  Self-Reflection: "The problem stems from the user's oversight of requirements, not the repository's deployment process. Keeping classification as question." |
| Round 3: Issue Raiser: "After upgrading Python, it works now. Thanks!" | Role Analysis: Issue Raiser confirms the solution after following Member's guidance.  Classification: question.  Evidence: "Fixing the Python version resolved the issue without changes to the repository."  Self-Reflection: "This confirms the issue was due to a misunderstanding of the README. Final classification should remain question." |

仅去除角色分析的Issue对话框架对Issue\_1的分类过程

|  |  |  |
| --- | --- | --- |
| Round | Input | Output |
| 1 | Training on a custom dataset produces low accuracy compared to the benchmark results. | Classification: performance Evidence: 'The issue appears to be related to poor model performance.' Self-Reflection: 'This could be due to dataset-related differences, but the issue still impacts performance.' |
| 2 | Tried augmenting the data, but accuracy remains low. | Classification: performance Evidence: 'Data augmentation did not help, so the issue might still be linked to performance.' Self-Reflection: 'The dataset differences might be more fundamental, but performance is still affected.' |
| 3 | Realized my dataset labels don’t match the expected format; correcting them improves accuracy significantly. | Classification: performance Evidence: 'After correcting labels, accuracy improved, but the initial classification remained performance-focused.' Self-Reflection: 'The issue stems from dataset misuse, but it was categorized as performance.' |

完整Issue对话框架对Issue\_1的分类过程

|  |  |  |
| --- | --- | --- |
| Round | Input | Output |
| 1 | Issue Raiser: 'Training on a custom dataset produces low accuracy compared to the benchmark results.' | Classification: question Evidence: 'Custom dataset issues often stem from format or pre-processing differences.' Role Analysis: 'Issue Raiser reports a dataset-related issue.' Self-Reflection: 'Could this be due to user dataset misuse?' |
| 2 | Member: 'Does your dataset format match the expected input as specified in the documentation?' | Classification: question Evidence: 'The Member highlights a potential mismatch in dataset format.' Role Analysis: 'Member identifies the likely source of the issue.' Self-Reflection: 'This reinforces that the problem is user-specific, not model performance.' |
| 3 | Issue Raiser: 'Realized my dataset labels don’t match the expected format; correcting them improves accuracy significantly.' | Classification: question Evidence: 'Correcting dataset labels resolved the issue caused by user-side configuration.' Role Analysis: 'Issue Raiser confirms it was a dataset issue.' Self-Reflection: 'This was a user misunderstanding, not a performance problem.' |

仅去除自我反思的Issue对话框架对Issue\_2的分类过程

|  |  |  |
| --- | --- | --- |
| Round | Input | Output |
| 1 | Issue Raiser: 'Training produces NaN values after a few iterations.' | Classification: error Evidence: 'NaN values indicate a potential bug in the repository’s training pipeline.' |
| 2 | Member: 'Are you using the recommended learning rate? The default value might be too high.' | Classification: error Evidence: 'The Member suggests tweaking a parameter, but the issue remains classified as a potential bug.' |
| 3 | Issue Raiser: 'Reducing the learning rate seems to fix the problem.' | Classification: error Evidence: 'Learning rate adjustments solved the issue, but the problem is still treated as an inherent bug.' |

完整Issue对话框架对Issue\_1的分类过程

|  |  |  |
| --- | --- | --- |
| Round | Input | Output |
| 1 | Issue Raiser: 'Training produces NaN values after a few iterations.' | Classification: question Evidence: 'NaN values could indicate an Issue with hyperparameter settings.' Self-Reflection: 'Could this be a usage problem rather than a repository bug?' |
| 2 | Member: 'Are you using the recommended learning rate? The default value might be too high.' | Classification: question Evidence: 'The Member’s suggestion points to incorrect user settings.' Self-Reflection: 'This reinforces the likelihood of a user misunderstanding rather than a bug.' |
| 3 | Issue Raiser: 'Reducing the learning rate seems to fix the problem.' | Classification: question Evidence: 'The Issue was resolved by following the Member’s advice.' Self-Reflection: 'The repository functions as expected. This was a user misunderstanding.' |