

AutoRL Assignment for AI Engineers Internship

Assignment: Implement a custom task using the AUTO RL component of OpenPipe/ART framework, demonstrating your understanding of RULER evaluation and GRPO training and train QWEN-0.6B model using this method.

The custom task is around picking the correct promotion for a given user. This means we need to return for one user which of the X promotions is best. We will share user profiles and the promotion profiles.

We do not provide any test data, please assume that gpt-4.1 verdict is correct.

Resources

- **Main Repository:** <https://github.com/OpenPipe/ART>
- **AutoRL Notebook:** https://github.com/OpenPipe/ART/blob/main/examples/auto_rl.ipynb
- **RULER Documentation:** <https://art.openpipe.ai/fundamentals/ruler>
- **RULER Blog Post:** <https://openpipe.ai/blog/ruler>
- **GRPO Explanation:** <https://www.oxen.ai/blog/why-grpo-is-important-and-how-it-works>

Setup

- Use Google Colab for development
- Use the OpenAI API key we provide
- Install: `!uv pip install -q openpipe-art==0.3.11.post2 langchain-core tenacity "gql<4" --prerelease allow`

Deliverables

1. **Demo** (video recording or live demonstration)
2. **Code** (send via email)
3. **Presentation** (maximum 5 slides explaining your approach and results)