AutoML-Agent

A Multi-Agent LLM Framework for Full-Pipeline AutoML

Patara Trirat[1] Wonyong Jeong[1] Sung Ju Hwang[1 2]

[1] DeepAuto.ai [2] KAIST

2025 ICML International Conference On Machine Learning

Project Page | OpenReview

🔍 Motivation

While AutoML significantly lowers the barrier to developing ML solutions, current frameworks:

- Require technical expertise to configure tools.
- Are often limited to specific pipeline steps.
- Do not leverage the full potential of LLMs.

🚀 Our Goal

Enable non-experts to develop end-to-end AI pipelines — from data retrieval to deployment — using only natural language descriptions.

💡 Key Contributions

✅ Full-pipeline AutoML using only natural language inputs.
✅ Task-agnostic and training-free pipeline planning.
✅ Combines retrieval-augmented planning, modular agent execution, and multi-stage verification.

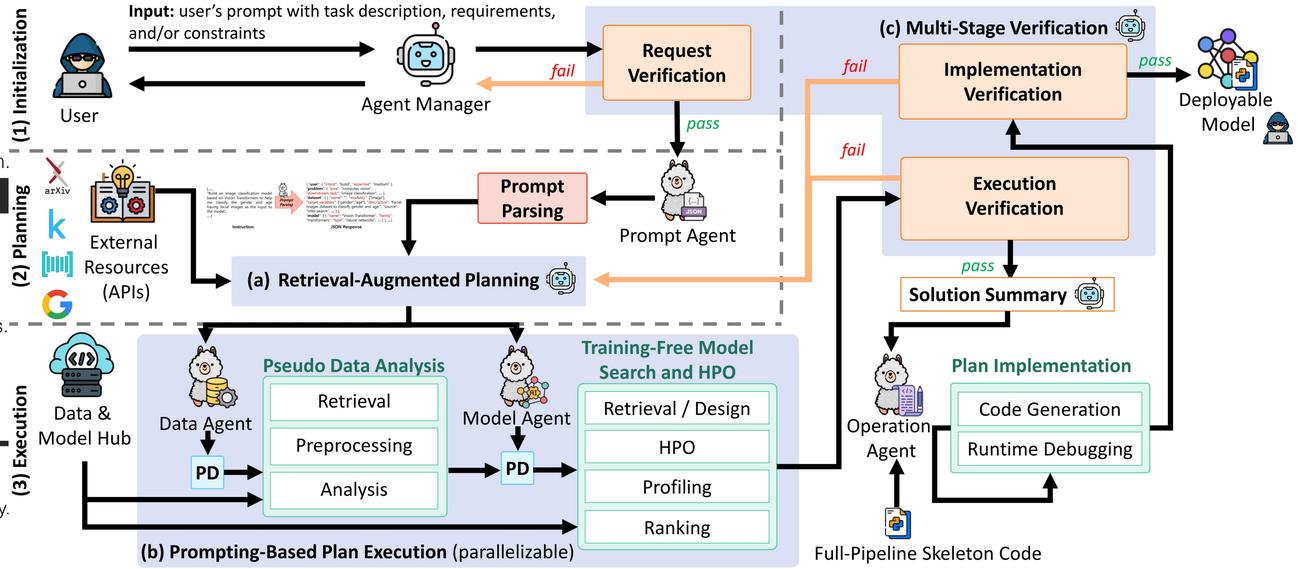
🏗️ Framework Overview

A multi-agent LLM framework that:

- Accepts user task instructions in plain language.
- Plans, executes, and verifies a complete ML pipeline.
- Delivers deployment-ready models with minimal human intervention.

Agent	Responsibility
Agent Manager	Orchestrates the workflow across all agents.
Prompt Agent	Parses user input into structured requirements.
Data Agent	Handles dataset retrieval, preprocessing, & analysis.
Model Agent	Conducts model search, HPO, and profiling.
Operation Agent	Implements, verifies, and deploys the final model.

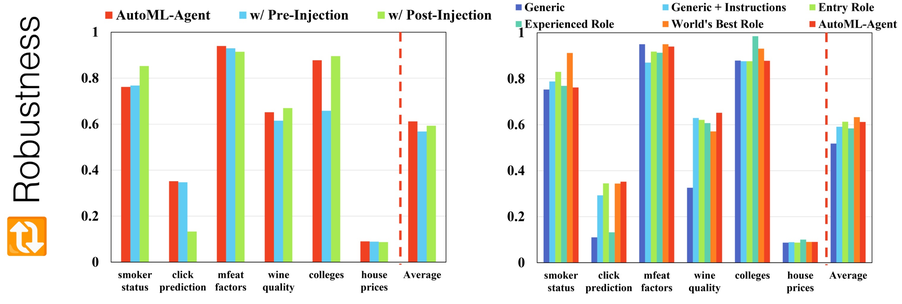
✔️ **Multi-stage verification** ensures quality at every step.
✔️ **Retrieval-Augmented Planning** enhances adaptability and efficiency.

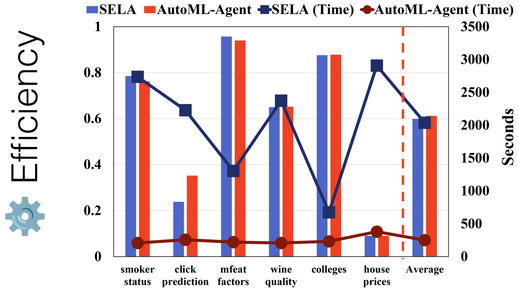


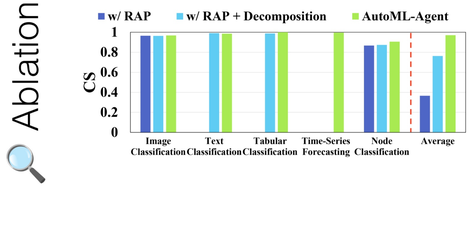
Overview Procedures

(1) Initialization stage aims to receive a valid user instruction using request verification.
(2) Planning stage focuses on extracting ML related information by parsing the user instruction into a standardized form (e..g, JSON), and uses it to devise plans accordingly.
(3) Execution stage executes each action given by the devised plans. Finally, based on the best execution results, AutoML-Agent outputs codes containing deployable model to the user.

📊 Experimental Results







Result Summary

- AutoML-Agent consistently ranks highest under both constraint-free and constraint-aware settings.
- Removing verification or planning components drastically reduces performance.
- Resilient to various prompts and noisy / misleading knowledge inputs.
- 8× faster than MCTS-based SELA framework with comparable or better accuracy.
- End-to-end deployment-ready pipeline costs ~$0.30 and 525 seconds.