

Date: Aug 12, 2025
To: "Vatsal Pravinbhai Patel" vatsal1@hawk-franklin-research.com;vatsal1804@gmail.com;vatsal.patel@iu-study.org;vatsal.patel@xaviers.edu.in
From: "Engineering Applications of Artificial Intelligence" support@elsevier.com
Subject: Your Submission EAAI-25-10859

Dear Mr. Vatsal Pravinbhai Patel,

The reviews of your manuscript: **EAAI-25-10859**; "Machine Learning Copilot Agent: An LLM-Guided Workflow for Prognostic Gene Discovery"; by Mr. Vatsal Pravinbhai Patel, Ananya Pal; Sourav Roy; Abhijeet Patel, MBBS, which you submitted to the International Scientific Journal Engineering Applications of Artificial Intelligence have now been received.

While the reviewers found the paper interesting, they are all agreed that very extensive revisions will be necessary if the paper is to be acceptable for publication in Engineering Applications of Artificial Intelligence, to make the paper of interest and benefit to a wider cross-section of our readers.

Please see below a copy of the suggestions of the reviewers, which need to be addressed before re-submission. I hope that you will find the reviewers remarks pertinent and helpful; we look forward to receiving a revised version of the paper. Please indicate, when submitting a revised version, what changes have been made to the paper. **Also, the revision of the paper should have no acronyms in the title and keywords, and all acronyms in the abstract must be defined.**

If you decide to revise the work, the revision due date is **Sep 02, 2025**.

To submit a revision, please go to <https://www.editorialmanager.com/eaai/> and log in with your username and password:

Your username is: *****

If you need to retrieve password details,
 please go to: *****

using 'Author Login'. Your paper can be found in "Submissions Needing Revision".

NOTE: Upon submitting your revised manuscript, please upload the source files for your article. We cannot accommodate PDF manuscript files for production purposes. We also ask that when submitting your revision, you follow the journal formatting guidelines. **Also, please note that the revised version should be formatted in single column.** For additional details regarding acceptable file formats, please refer to the Guide for Authors at: <https://www.elsevier.com/journals/engineering-applications-of-artificial-intelligence/0952-1976/guide-for-authors>

We sincerely appreciate having been given the opportunity to consider this manuscript.

Include interactive data visualizations in your publication and let your readers interact and engage more closely with your research. Follow the instructions here: <https://www.elsevier.com/authors/author-services/data-visualization> to find out about available data visualization options and how to include them with your article.

Research Elements (optional)

This journal encourages you to share research objects - including your raw data, methods, protocols, software, hardware and more – which support your original research article in a Research Elements journal. Research Elements are open access, multidisciplinary, peer-reviewed journals which make the objects associated with your research more discoverable, trustworthy and promote replicability and reproducibility. As open access journals, there may be an Article Publishing Charge if your paper is accepted for publication. Find out more about the Research Elements journals at <https://www.elsevier.com/authors/tools-and-resources/research-elements-journals>
 dgcid=ec_em_research_elements_email.

With very best wishes -

Patrick Siarry
 Editor In Chief,
 Engineering Applications of Artificial Intelligence

Reviewer comments:

Editor: All reviewers have provided comprehensive feedback on your paper. Here are my suggestions:

- 1 - Please revise the abstract so it is clearer and more concise. Please try to answer the following questions: a. What is the relevance of the study? (Motivation and Justification); b. What have you done? (Aim of the study); c. How have you done it? (Methodology); d. What have you found? (Findings); e. Why does it matter? (Significance and originality). [do not include the questions or category names in the revised abstract].
- 2 - Too many graphs are combined as Figure 2. This made them so small and unintelligible. Please divide Figure 2 into meanin 3 figures and only include the ones that are absolutely necessary.
- 3 - Please reconnect with the literature that informed your work in the discussion section.



I look forward to receiving your revision.

Reviewer #1: The authors present the Machine Learning Copilot Agent to automate ML workflows in precision oncology, specifically for non-

expert users lacking deep programming skills. The agent orchestrates a series of modular analytical tasks, including data cleaning, model training, feature selection, and survival analysis, within a secure, sandboxed environment. The authors showcase the agent through a case study in HNSCC.

Here come my comments:

1. The execution presented in the paper occurs in a sandboxed environment. However, it's less clear how the system deals with generated code failures, subtle bugs, or dataset anomalies (e.g., missing/malformed data). More detail on error correction, rollback strategies, and user notification in ambiguous cases would be instructive.
2. The agent-driven workflow is compared mainly to traditional scripting/manual pipelines. The authors may want to consider benchmarking against state-of-the-art AutoML and existing LLM-agent systems (e.g., AutoGen, Agno, HuggingGPT, LangChain agents), both in terms of accuracy and user effort required.
3. The authors mentioned about the importance of data privacy and compliance in Clinical ML tasks in section 1.1, but they didn't address the issue later in the paper at all.
4. ML-driven biomarker discovery is known to be vulnerable to overfitting, especially when testing many combinations of features and classifiers in modestly sized datasets. The paper does not provide detail on independent replication using truly out-of-distribution data, nor addresses the risks of selection bias that can arise from extensive feature/model exploration without holdout sets or blinded validation.

Minor issues:

1. Fonts in the figures are too small.
2. A couple of missing references to figures (e.g. Figure ??)

Reviewer #2: This paper presents the Machine Learning Copilot Agent, an AI-driven system that uses GPT-4 and LlamaIndex to automate complex biomedical machine learning workflows via natural language commands. Through a case study on prognostic gene discovery in Head and Neck Squamous Cell Carcinoma (HNSCC), the agent executed an end-to-end pipeline—including clustering, classification, and survival analysis—without manual coding, successfully identifying a robust gene signature validated across two patient cohorts.

While the system demonstrates technical feasibility and usability, several limitations remain to be addressed below:

- 1) While the integration of LLMs into machine learning workflows is novel from a systems perspective, the bioinformatics and ML methodology itself is standard. Clustering (KMeans, GMM), classification (RF, SVM, LR), and feature selection (SelectKBest) are all established techniques. The novelty lies in the orchestration of these steps using LLMs, not in the analytical methods.
- 2) The system relies heavily on custom instructions, which requires users to input detailed and precise natural language prompts. This shifts complexity from programming to prompt engineering, potentially limiting usability for non-experts.
- 3) The analysis focuses on how the system executed tasks, not on the biological interpretation of results. There is little discussion about biological plausibility or clinical relevance of the identified gene signatures.
- 4) The agent's performance is only tested on one case study (HNSCC gene expression data). No demonstration across other cancer types, ML tasks (e.g., regression, unsupervised learning outside clustering), or data modalities (e.g., images, proteomics). As such, the system's domain generalizability is unproven.
- 5) The model performance is evaluated using AUC and log-rank p-values, which are helpful but insufficient alone. No metrics for model calibration, precision-recall, or clinical decision curve analysis.
- 6) The results alternately mention a 15-gene signature (abstract) and a 12-gene signature (results section), then mention 200 genes (discussion). This inconsistency undermines the clarity and reliability of the key findings.
- 7) No direct benchmark comparison with alternative automation platforms (e.g., AutoML tools, HuggingGPT, LangChain, Google Cloud AutoML). Lacks quantitative evaluation of speed, error rate, or efficiency gains compared to manual scripting.

Reviewer #3: This manuscript presents the Machine Learning Copilot Agent, a novel LLM-guided framework designed to automate complex biomedical machine learning workflows through natural language interaction. The authors showcase a compelling case study in precision oncology, demonstrating the agent's ability to identify robust prognostic gene signatures in head and neck squamous cell carcinoma using TCGA and CPTAC datasets. The paper clearly outlines the system's architecture, operational flow, and performance, highlighting the practical potential of LLM-powered agents in complex scientific domains. Overall, the manuscript is well-written and presents a technically interesting and timely study. The case study is thorough and well validated, and the proposed agent system demonstrates real-world applicability. I recommend acceptance after minor revision.

(1) There are a few minor typographical errors in the manuscript. For instance, on page 3, the sentence "A conceptual diagram of the architecture is shown in Figure ??" appears to be incomplete or contains placeholder symbols. Please revise such instances.

(2) The readability of Figure 1 can be improved. The font size is quite small, and the color contrast in some parts makes interpretation difficult. Consider enlarging labels and optimizing the color palette for better clarity, especially for print readers.

(3) While the manuscript presents a solid overview of related work, the literature review could be further strengthened by incorporating a broader discussion of recent developments in simulation research automation and machine learning workflow automation. In particular, several recent studies have proposed frameworks that enable end-to-end automation of scientific analysis tasks based on high-level research goals, without requiring manual coding or user intervention. For example, one JCIM work introduced a brief-rp system designed for simulation automation, which can interpret simple research objectives and autonomously generate and execute multiple machine learning pipelines to complete the analysis. Discussing such approaches would help contextualize the novelty and scope of the ML Copilot Agent, especially in terms of automation granularity and user interaction models.

(4) The manuscript inconsistently refers to different feature set sizes throughout various sections, which may cause confusion for readers. For example, the abstract and main text sometimes mention a 15-gene signature, elsewhere a 12-gene signature is described as the final model, and Section 5 refers to a 200-gene signature derived from Logistic Regression. While these may reflect different stages in model development (e.g., exploratory vs. final models), the relationships between them are not clearly explained.

3

Reviewer #4: This manuscript presents the "Machine Learning Copilot Agent," an LLM based workflow automation framework, demonstrated through a multi-cohort prognostic biomarker discovery case study in Head and Neck Squamous Cell Carcinoma (HNSCC). The technical integration of GPT-4o with LlamaIndex, combined with secure sandboxed code execution, is an interesting contribution to reproducibility and usability in data-intensive biomedical research.

However, the manuscript requires substantial revision before it could be considered for publication in Engineering Applications of Artificial Intelligence. My major concerns are outlined below

Title/Abstract Misalignment: The title is broad and generic ("Prognostic Gene Discovery") and does not reflect the specific case study focus or the main technical novelty. The abstract is narrowly biomedical, with minimal emphasis on cross-domain applicability, and it repeats the result inconsistency noted above. Both should be rewritten to clearly match the scope, contributions, and final outcome.

Language, Formatting, and Completeness:

Several typographical errors ("hand's on coding," "re-classified") and grammatical issues require correction.

Some figure captions contain placeholders ("Figure ??") and should be finalized.

References contain placeholders (e.g., "zenodo.1234") and incomplete metadata.

Result:

The manuscript alternately presents a 15-gene, 12-gene, and 200-gene "final signature" as the main finding. These conflicting figures appear in the abstract, results, and discussion, without a clear explanation of why they differ or which is definitive. If multiple "best" signatures exist under different evaluation criteria, this should be explicitly stated and framed accordingly.

General:

The workflow is described in detail multiple times in Sections 3, 4, and again in the Appendix often with only minor changes in wording.

Ethical/data privacy compliance should be explicitly stated for the biomedical datasets used, confirming that they are publicly available and anonymized.

The manuscript has good potential to make a meaningful contribution, particularly at the intersection of LLM and applied engineering research. However, for better adaptability across diverse domains and to sustain the interest of readers. The suggested changes would be valuable in enhancing the manuscript's clarity, consistency, and overall appeal.

No acronyms may be used in the title.

No acronyms may be used in the keywords.

Acronyms in the abstract must be defined on first usage.

Both the implemented AI, as well as the application of AI must be mentioned in the keywords and abstract

These rules are applicable to all acronyms. This includes commonly used acronyms (e.g. AI, ML, 3D), model names (e.g. YOLO, swin transformer, BERT), dataset names (e.g. KITTI, CEC2017, PASCAL VOC), units (e.g. mm, MPa, GFlops), algorithm names (e.g. PSO, LASSO), and novel proposals (e.g. names of new architectures)

The title, keywords, and abstract entered on the EM system must match the corresponding entries in the manuscript; care should be taken to ensure that no equation references appear incorrectly on the EM system.

For further assistance, please visit our customer support site at <http://help.elsevier.com/app/answers/list/p/7923>. Here you can search for solutions on a range of topics, find answers to frequently asked questions and learn more about EM via interactive tutorials. You will also find our 24/7 support contact details should you need any further assistance from one of our customer support representatives

At Elsevier, we want to help all our authors to stay safe when publishing. Please be aware of fraudulent messages requesting money in return for the publication of your paper. If you are publishing open access with Elsevier, bear in mind that we will never request payment before the paper has been accepted. We have prepared some guidelines (<https://www.elsevier.com/connect/authors-update/seven-top-tips-on-stopping-apc-scams>) that you may find helpful, including a short video on Identifying fake acceptance letters (<https://www.youtube.com/watch?v=o5l8thD9XtE>). Please remember that you can contact Elsevier's Researcher Support team (<https://service.elsevier.com/app/home/supporthub/publishing/>) at any time if you have questions about your manuscript, and you can log into Editorial Manager to check the status of your manuscript (https://service.elsevier.com/app/answers/detail/a_id/29155/c/10530/supporthub/publishing/kw/status/).

#AU_EAAI#

To ensure this email reaches the intended recipient, please do not delete the above code

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. ([Remove my information/details](#)). Please contact the publication office if you have any questions.