IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

Self-Healing and Self-Learning AI Systems: A Framework for Error Detection, Correction and Memory Retention in Generative AI Pipelines

Aniket Dilip Bhujbal[1], Kashish[2], Radhika Mantri[3]
Student, Department of Mathematics
Fergusson College (Autonomous) Pune , India[1,2,3]

Abstract: *Recently, there has been a huge rise in the number of Generative AI systems being used throughout our world for tasks such as Decision Making, Automation and Content Generation. Unfortunately, many Generative AI systems do not produce consistent or correct outputs which can degrade their overall credibility. A major shortcoming of current AI systems is their inability to recognize their own mistakes and correct them. Therefore, similar errors could be repeatedly created, especially in real-world uses, where accuracy is a major concern.*

In this paper, we provide a Self-Healing and Self-Learning AI framework that will detect errors, correct them and retain a record of prior mistakes. Our Approach consists of three Phases. The first Phase will analyze the generated output from Generative AI to determine if there are any potential errors in it using basic Validation and Feedback methods. The second Phase will either refine or regenerate the response in order to create a more Accurate and more Relevant result.

The last feature of the learning loop is the memory function of storing past errors along with their corrections to allow the system to learn from previous errors.

The learning loop creates a recursive feedback loop in the AI pipeline that allows the system to build on its errors to become more efficient at making correct decisions over time. The primary goal of this proposed method is to produce a more reliable, adaptable and efficient AI system for use in critical instances where accurate decision making is very vital.

Keywords: Generative AI, Self-Healing Systems, Error Detection, Feedback Loop, Adaptive Systems, AI Pipelines

I. INTRODUCTION

In the past few years, Generative Artificial Intelligence (AI) has become one of the most revolutionary technologies found today with multiple uses including chatbots, recommendation systems, content generation and decision support systems. Despite its rapid growth, Generative AI continues to be challenged by issues of reliability, accuracy and consistency in its output. Many times, the response are generated incompletely or incorrectly which limits the amount of trust that users can have in these Generative AI systems.[1]

Current generative AI model limitations restrict them all from being able to evaluate their own outputs to learn from prior mistakes. Generally speaking, these generative AI systems create outputs but do not self-evaluate or learn as a result of generating incorrect output. In situations such as finance, healthcare and cybersecurity, producing incorrect generative AI outputs could result in dire consequences.

The growing need for Adaptive AI Systems will permit these generative AI systems to constantly monitor, evaluate and improve their performance. In this paper, we present a Self-Healing and Self-Learning AI Framework based on error detection, error correction and long-term memory so that these generative AI systems can become more reliable and adaptable.[3]

DOI: 10.48175/IJARSCT-36853

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

II. LITERATURE REVIEW

New developments in generative AI have produced new architectures and new training methodologies that improved how well models respond to user requests. Examples include Reinforcement Learning from Human Feedback (RLHF) and Retrieval-Augmented Generation (RAG). Both types of techniques have enhanced the quality of responses provided by generative AI models; however, they are highly reliant on the availability of external information or retraining and therefore do not enable continuous autonomous real-time corrections. [2]

While some research has examined adaptive systems and feedback-based learning, most of these systems use different methodologies and would require separate implementations. Most AI systems currently do not have a unified framework that enables real-time detection and correction of errors or long-term memory storage of experiences. As a result, many of the available systems would continue to produce the same types of errors due to the current state of design and implementations in AI systems. To remedy this issue, we propose a framework to facilitate continuous self-improvement of generative AI pipelines.[5]

In recent studies on self-healing software, there has been a significant focus on the integration of detection, diagnosis, and recovery into a single system. This has been done in an attempt to minimize human intervention as much as possible.

While there have been significant developments in this field, most existing techniques have only been able to deal with individual components. Not much research has been done on how to bring all these components together into a single system that allows for learning.[3]

This paper, therefore, proposes a unified system in artificial intelligence that allows for self-healing and self-learning through error detection, correction, and retention.

III. PROBLEM STATEMENT

There are several limitations with the functioning of current generative AI systems that have made them less effective for use in real-world applications. For example, generative AI systems do not have the ability to identify incorrect outputs independently, nor can they automatically correct an output once they have detected an error. Moreover, they do not have any means by which to keep track of their previous errors; thus generating repeated errors over a period of time.

The lack of self-awareness and adaptability of current generative AI systems results in diminished reliability of these systems when used in applications requiring repetitive, accurate, and consistent decision-making. As a result, there is a need to establish an AI framework that supports the continuous monitoring of its outputs to allow it to automatically correct errors due to incorrect outputs, as well as learn from its historical output data.

IV. PROPOSED & FRAMEWORK

The proposed framework will add an element of self-healing and self-learning to the generative AI pipeline. The basic elements of this framework are three interrelated components that will all function cohesively to enhance performance of the overall system.[5]

The first component is called the error detection module which evaluates the generated outputs against validation rules, the confidence score, and by conducting contextual analysis. By using these methods, the error detection module will determine if the output is correct, unsure or incorrect.

The second component is the correction module. The correction module is a module that triggers when the error detection module detects an error in the output generated by the system. The correction module will either refine or regenerate the response by changing the prompt used to generate the response, providing additional context for the generation of the output or by re-processing the input through the system to enhance the quality of the output generated.

The third component is the memory module. The memory module is responsible for storing data that pertains to previous errors and the corrections made to correct those errors, as well as to contextual information about the previous

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

errors. This enables the generative AI system to learn from prior mistakes and avoid making these mistakes when generating subsequent outputs.

These three interrelated components of the system will create an ongoing feedback loop allowing the generative AI system to improve and evolve over time.

V. METHOLODOGY

The proposed system will follow a structure in which a multi-level AI system will be used, incorporating a language model generator, error detection, error correction, and error memory. In this case, the proposed system will incorporate a language model generator, which will be the Large Language Model (LLM). However, it is worth noting that a pre-trained LLM will be used in the base layer of all the proposed generative responses.[6]

In this case, the LLM will create a response depending on the input, which will then be forwarded to the Error Detection stage, which will be validated depending on the Rule-based validation, keyword validation, confidence, etc. However, this will contain an error in case there is an inconsistency in the response, which will also be of low quality. However, in case a problem arises, the Correction module will be invoked, which has the possibility of changing the prompt, re-querying the LLM, and/or utilizing the context information in order to generate a good output. There is a possibility that a few iterations might take place in order to generate a proper output.

There is a memory module in place, and the memory module has stored all the previous queries. Therefore, in case a problem arises, the output of the problem will also be stored. This memory module would improve the output and would not allow the same error to take place once again in case the same query is made once again.

This is a loop, and this would improve over a period of time based on the queries. This framework has to be tested, and the testing of the entire framework is based on the accuracy of the output, reduction of repeated errors, and qualitative improvements in the output.

VI. ARCHITECTURE

In summary, a modular architecture will serve as the foundation for the proposed system architecture which will continually monitor, improve, and correct AI content created by the proposed system architecture; the proposed architecture will be established through building the AI model and providing validation and correction components; the components of the proposed system architecture include:

- Input Layer,
- Generative Model,
- Error Finding Module,
- Corrections Module,
- Memory

Users will provide input to the API. The query will be processed using the generative model within the AI. For this generative model, pre-trained models - transformer, i.e. GPT - will be used to develop generative models of the AI because pre-training provides an understanding of a natural language.[2]

After the response is generated, it is not sent back to the user. Rather, it is sent to the error detection module. This module is of vital importance in determining if the generated response is correct, relevant, and comprehensive. This is carried out through simple validation mechanisms. Furthermore, the relevance of the generated response is checked through a comparison of the generated response with the original query. If the generated response is found to be incorrect or of low confidence, it is sent to the correction module. This module refines the generated response through various mechanisms, such as refining the prompts, re-generating the output, and adding more context. In some cases, the process is repeated several times until the desired quality is achieved.[11]

The final module is the memory module. This module stores the previously interacted information, which includes the identified errors and the corresponding corrected responses. This module can be easily implemented through simple mechanisms, such as databases or vector-based memorySystem Flow

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

1. From user's point of the view, the system works as Follows:
2. User enters query into the system.
3. The generative AI model produces an initial response.
4. The system checks the response for errors or inconsistencies
5. .If an error is detected, the response in refined automatically.
6. The improved response is generated and shown to the user.
7. The system stores the mistake and correction for future learning.

VII. RESULT & APPLICATION

From the above-proposed framework, it is clear that the accuracy, consistency, and reliability of the output generated from the AI system have increased. The main reason for this is the inclusion of the error detection and correction feature, which minimizes the chances of incorrect output generation from the AI system. This has further been enhanced through the inclusion of the memory module, which minimizes the chances of repetition of the same error.

If the above-proposed self-healing framework is compared to the conventional generative AI systems, it is clear that the proposed AI system has a lot of potential to be adapted and enhanced. Though the results may vary, the concept has a lot ofpotential.

VIII. EVALUATION METRICS

Several metrics are used to assess the self-healing and self-learning AI framework's performance. Since the system is proposed, the evaluation's foundation is comprised solely of anticipated advancements and theoretical analysis.[8]

The first metric, the accuracy improvement rate, is a metric for determining the extent to which the system will be able to improve the overall correctness of response generation through the enhancement process of the output generated by the response generation algorithm. Therefore, it is expected that the refinement process will lead to the reduction of erroneous and incomplete output.

The second metric, error recurrence rate, assesses the system's ability to mitigate the frequency of error occurrence due to the system using a memory module to keep track of previous errors. As a result, when the system develops its first incorrect response to a certain question, it will remember this incorrect response and reproduce it if a similar question is asked in the future.

Thirdly, response consistency refers to the extent to which the output of the system continues to produce stable and reliable responses to a similar question. Because the self-healing and self-learning AI framework contains a continuous feedback loop that provides feedback to itself, the expected consistency of its output will improve over time at an increasing rate.

Lastly, the quality of responses can also be determined based on the following criteria: relevance; completeness; clarity. The refining of the original output will enhance each of these characteristics due to the correction mechanism implemented within the framework. [10]

A conceptual comparison between traditional AI systems and the proposed framework is shown below:

❖ **Metric**	❖ **Traditional AI**	❖ **Proposed System**
❖ **Accuracy**	❖ Moderate	❖ Expected High
❖ **Error Repetition**	❖ High	❖ Expected Low
❖ **Consistency**	❖ Low	❖ Expected High
❖ **Adaptability**	❖ Low	❖ Expected High

The results of this research indicate that the proposed framework could be beneficial for improving generative AI systems regarding both efficiency and dependability.[7]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, June 2026
Impact Factor: 8.2

VIII. APPLICATION

The proposed framework can be applied to different areas where there is a need for reliability and accuracy.In the healthcare sector, the framework can be applied to assist in medical suggestions without the possibility of incorrect information being provided. This is ensured through the self-healing mechanism.In finance, the framework can be applied to decision-making systems to eliminate errors in prediction and recommendation. This will assist in the minimization of financial risks.[3]In education, the framework can be applied to learning systems to assist in improving the quality of explanations provided to the user.In customer support systems, the framework can be applied to assist in improving the quality of performance of the chatbots.The framework that is proposed has a wide range of applications, including intelligent chatbots, financial advisory systems, decision support systems in the medical field, educational platforms, and customer service automation. All of these require greater accuracy and reliability for effective functionality.[5]

X. FUTURE SCOPE

The framework has much potential for further improvements in the future. For example, there is the possibility of incorporating real-time data processing into the framework.

There is also the possibility of improving the framework through the use of advanced learning techniques such as reinforcement learning. This will further enable the framework to learn from interactions.

The framework is also likely to be extended to the Internet of Things (IoT), where there is a need for real-time monitoring of the system as well as recovery from faults. In such a case, the self-healing mechanism will be beneficial. [7]

Possible areas for future research include integrating sophisticated learning algorithms, improving memory storage techniques, and integrating real-time data streams. Moreover, improving evaluation metrics for detecting errors will also improve performance.

XI. CONCLUSION

Other than this, the proposed framework will also focus on the need to make the AI system more adaptive and intelligent in nature. It is to be noted that, unlike the conventional generative AI system, which is more static in nature, the proposed system is more dynamic in nature as it will allow the system to assess the quality of output being generated.

In this paper, a framework for a self-healing and self-learning AI system will be presented. The main objective of this framework will be to enhance the learning capabilities of generative AI systems. The framework will be based on various components of error recognition, correction, and retention. This will help the AI system learn to improve its performance gradually. This is a major breakthrough in the creation of intelligent AI systems with learning capabilities.

REFERENCES

1. Adebayo, Y., et al. (2026). *Predictive failure detection and self-healing CI/CD pipelines using AI: A framework for intelligent software delivery*. ResearchGate. https://www.researchgate.net/publication/400668320_Predictive_Failure_Detection_and_Self-Healing_CICD_Pipelines_Using_AI_A_Framework_for_Intelligent_Software_Delivery

2. Rivers, J. A., Brooks, N. M., Owens, C. J., Harrington, M. T., & Song, M. (2026). *Securing generative AI pipelines in finance*.ResearchGate.https://www.researchgate.net/publication/400509929_Securing_Generative_AI_Pipelines_in_Finance

3. Akhtar, M. H., Ghafoor, U., Imran, O., Ayub, N., Abdullah, M. M., & Khan, H. (2026). *An efficient AI and deep learning assisted self-healing network approach: Analysis on fault detection, response and recovery to mitigate threats in IoT-security ecosystem*.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 7, June 2026 Impact Factor: 8.2

https://www.researchgate.net/publication/400446987_An_Efficient_AI_and_Deep_learning_Assisted_Self-Healing_Network_Approach_Analysis_on_Fault_Detection_Response_and_Recovery_to_Mitigate_Threats_in_IoT-Security_Ecosystem

4. Kalakoti, M. K. R. (2025). *AI-augmented self-healing infrastructure: Combining health probes with remediation playbooks*. *Journal of Information Systems Engineering and Management, 10*(58s) https://www.researchgate.net/publication/395424309_AI-Augmented_Self-Healing_Infrastructure_Combining_Health_Probes_with_Remediation_Playbooks

5. Sawalkar, V., Jagadale, S., Shendkar, B., Chandre, P., & Mhaske, C. (2025). *Self-healing cloud infrastructure: Leveraging AI for fault detection and recovery*. In *Data Science and Big Data Analytics*. Springer. https://www.researchgate.net/publication/396078460_Self-Healing_Cloud_Infrastructure_Leveraging_AI_for_Fault_Detection_and_Recovery

6.Kalakoti, M. K. R. (2025). *AI-augmented self-healing infrastructure: Combining health probes with remediation playbooks*. *Journal of Information Systems Engineering and Management, 10*(58s), 1147–1157. https://www.researchgate.net/publication/395424309_AI-Augmented_Self-Healing_Infrastructure_Combining_Health_Probes_with_Remediation_Playbooks

7. Bradfield, E., Woo, K., & Adebayo, Y. (2025). *Generative AI for pipeline repair: Automating corrective actions in self-healing data systems*. ResearchGate. https://www.researchgate.net/profile/Yusuf-Adebayo-2/publication/396161486_GENERATIVE_AI_FOR_PIPELINE_REPAIR_AUTOMATING_CORRECTIVE_ACTIONS_IN_SELF-_HEALING_DATA_SYSTEMS/links/68dfe672d221a404b2a531ac/GENERATIVE-AI-FOR-PIPELINE-REPAIR-AUTOMATING-CORRECTIVE-ACTIONS-IN-SELF-HEALING-DATA-SYSTEMS.pdf

8. James, M. (2025). *Future of self-healing machines: Autonomous maintenance enabled by generative AI*. ResearchGate. https://www.researchgate.net/publication/395917946_Future_of_Self-Healing_Machines_Autonomous_Maintenance_Enabled_by_Generative_AI

9. Peck, K. (2025). *Reconceptualizing pedagogical frameworks for the AI epoch: Advancing critical and reflective thinking in generative AI-mediated learning environments*. ResearchGate. https://www.researchgate.net/publication/399987071_Reconceptualizing_Pedagogical_Frameworks_for_the_AI_Epoch_Advancing_Critical_and_Reflective_Thinking_in_Generative_AI-Mediated_Learning_Environments

10. Sarma, S. S., Gami, S., & Balogun, P. (2025). *Self-healing software: AI-driven automated debugging and error recovery mechanisms*. ResearchGate. https://www.researchgate.net/publication/396823641_Self-Healing_Software_AI-Driven_Automated_Debugging_and_Error_Recovery_Mechanisms_SUBRAMANYA_SARMA_S_SANDIP_GAMI

11. Watson, S. (2025). *Alignment of the starlings: Learning with generative AI*. *Future Humanities*. https://www.researchgate.net/publication/399139146_Alignment_of_the_Starlings_Learning_With_Generative_AI

12. Rauba, P., Seedat, N., Kacprzyk , K., & van der Schaar, M. (2024). *Self-healing machine learning: A framework for autonomous adaptation in real-world environments*. arXiv. https://arxiv.org/pdf/2411.00186

13 . Manne, T. A. K. (2024). *Generative AI for cloud infrastructure decision-making and self-healing systems*. *Journal of Artificial Intelligence & Cloud Computing, 3*(3), 1–5. https://www.onlinescientificsresearch.com/articles/generative-ai-for-cloud-infrastructure-decisionmaking-and-self-healing-systems.pdf