The Poetry of Error: Analyzing Mistakes in AI-Generated English Works

Humera Nafees[1], Mohammed Kaif[2], Mohammed Afnan Khan[3], Abdul Haseeb[4], P Kaif Ali Khan[5]
[1]*Assistant Professor Lords Institute of Engineering and Technology- Hyderabad*
[2,3,4,5] *BE 1st Year Lords Institute of Engineering and Technology- Hyderabad*

Abstract- **Artificial intelligence is playing a big role in creative writing, with tools like Deepseek and ChatGpt generating stories, poems, and articles that can sound surprisingly natural. But if we observe them closely, there's still something missing. These AI-generated texts often seems to be emotionally flat, sometimes awkward, and also sometimes repetitive. These AI Models struggle with things like keeping a consistent tone, understanding deeper meanings behind words, something we usually do without thinking twice.**

Our paper, "The Poetry of Errors: Analyzing Mistakes in AI-Generated English Works," looks at those common errors. By going through samples of AI-written poems and stories, it points out where AI tends to mess up and whether it's misusing literary devices, missing emotional cues, or just sounding too mechanical.

Our paper doesn't stop at the problems, though. This paper also suggests many ways to fix or at least improve these issues. The solution includes training AI Models with more context-rich data which has emotions, designing systems that learn from human feedback along with their mood analysis, and encouraging human-AI collaboration instead of relying on these AI Models completely. Adding better cultural understanding and real-time editing by people will also help to solve this problem. In the end, even though AI is a great tool for writers, this paper argues that the human touch which has our creativity, emotions, and instinct is still something that algorithms and their few lines of syntax can't fully copy.

I. INTRODUCTION

Artificial Intelligence (AI) has quickly made it's way from the sidelines to the center of modern innovation. What we used to think to be science fiction in the past is now a part of our daily reality, whether it's using voice assistants to set reminders, getting personalized suggestions from streaming platforms, seeing automatic vehicles navigate traffic like Tesla. We all can see AI's influence everywhere around us, but one of its most surprising and debated roles has emerged in the field of creative writing and English Literature.

1.1 AI models like OpenAI and Deepseek have demonstrated an amazing ability to produce English literature from poems and short stories to reflective essays, novels and plays. These tools can mimic human language with such fluency that, at first glance, it can be hard to tell whether a passage was written by a person or a machine. This sudden creative capability has caught the attention of writers, educators, and technologists alike, sparking both enthusiasm and concern.

1.2 As we know that AI is great tool for English Literature, but it is also very important to understand where it is lacking. Creative writing, especially in English literature is not just putting words together correctly. It's about turning our imagination into language, creating rhythm, setting the right tone and expressing deep emotions. These are some of the areas where AI usually struggles. The writing seems polished in the beginning, but it feels flat underneath, missing the emotional depth and originality that defines true human writing.

1.3 In our paper, we look at those problems in detail. We look at the most common mistakes AI comits when trying to write English literary pieces. These include weak themes, sudden and awkward changes in tone, repetitive phrases, emotionally off-key expressions, and shallow structure. By studying AI-generated poems, stories, and essays, we try to understand not just what these mistakes are, but why they happen—and what they tell us about the writing process itself.

1.4 The title "The Poetry of Error" signifies a lot of meanings. It throws light to the obvious mistakes found in AI-generated literature at the surface. But it also suggests something deeper, that even in its

flaws, AI reveals something meaningful. Those mistakes show us what makes our writing, the human writing truly special and remind us of the subtle skills we often don't even realize we use like emotional intuition, personal memory and creative intent.

1.5 This paper does more than point out flaws but it also looks ahead. We explore practical solutions to improve the quality and originality of AI-generated literary content. These include developing systems with stronger and better emotional understanding, adding and upgrading tools for better contextual awareness and encouraging more interaction between human writers and AI models. Our goal should be to use AI to improve and enhance our works rather than entirely being dependent on them.

1.6 Looking at the bigger picture, this research joins an important and timely conversation: what should AI's role be in the future of English literature? As AI tools become more common and easier to use, it's not enough to look at just what they can do but we also need to think about what they can't do and why human creativity still matters so much at this time where technology is taking over everything. By noticing the small mistakes that AI is making, we actually learn more about our own abilities as writers. These errors remind us what makes human storytelling special, and they can help us find new ways to work together with technology as the world of literature continues to grow and change.

II. LITERATURE REVIEW

2.1 Artificial intelligence has made its way into the field of creative writing with a lot of force over within the last few decades. Mass datasets including novels, poetry, essays, and a broad range of written materials are used to train AI models. Often echoing the tone and style of human writers, these tools can generate material that seems coherent and grammatically accurate. Researchers such as Floridi and Chiriatti (2020) have questioned the authenticity of machine-generated creativity by investigating how these models function on probability-based prediction instead of real understanding.

2.2 Research by Kotek and Rudy (2021) highlight how frequently AI-generated books lack emotional realism and contextual depth. Though on a surface level the language may sound convincing, artificial intelligence usually lacks the subtlety, cultural context, and nuance human literature demands. Mahowald and Fedorenko (2022) claim that one of the main shortcomings of artificial intelligence in literature is its incapacity to fully capture emotional tone, so producing uneven or shallow representations in poetry and prose.

2.3 The overuse of cliches, repetitious phrases, and structurally awkward outputs in AI-written content has been noted by several researchers (Todorov, 2021; McNeill, 2023). These problems arise because AI can only remix preexisting patterns in the texts it was trained on. AI does not have emotional memory and lived experiences. The result of this research was that AI can mimic and fake feelings but can hardly ever signify them deeply.

2.4 Nonetheless, some research emphasizes AI's potential as a co-creator. According to Jakesch (2019), we, the human authors and AI can work together to create works that take advantage of both human creativity and machine's efficiency. This cooperative strategy is thought to be a workable way to include AI into creative industries without displacing human authors.

2.5 A thorough examination of the precise kinds of mistakes AI makes in literary content—in particular, emotional mismatch, thematic flatness, tonal inconsistency, and unnatural phrasing—is still missing from a large portion of the body of current literature. While acknowledging the limitations, the majority of studies do not provide a thorough analysis or specific solutions to enhance AI's performance in literature.

2.6 By analyzing the recurrent problems in AI-generated English literary content and offering workable solutions, this paper seeks to close that gap. Enhancing AI systems' emotional intelligence by employing prompt engineers with English background for specifically this purpose, encouraging cooperative human-AI editing processes, and adding feedback mechanisms to improve output quality are a few of these. By doing this, we intend to close the gap between human literary artistry and artificial generation.

III. METHODOLOGY

Our study examined the various types of mistakes artificial intelligence usually makes while writing

English literature and we also suggest important fixes using a technique called qualitative content analysis method. The study examined literary works produced by AI, primarily poems, short stories, and narrative essays, using GPT-3 and GPT-4, two of the most advanced language models available today.

1. Data Collection

We gathered literature works using prompts which had a variety of emotions, themes and genres. We asked prompts such as:

i)"Compose a poem for me about loss."
ii)"Write a brief narrative about friendship."
iii)"Write an emotional essay describing a personal loss."

In order to preserve authenticity and replicate normal user interaction, the texts were produced with little assistance from humans using OpenAI's ChatGPT interface.

2. Error Identification

We reviewed each AI-generated work carefully and annotated them:

i)Emotional flaws: There was an issue of emotion, forced sentiment and emotionally inappropriate expressions.
ii)Tone inconsistencies: Sudden or unnatural shifts in voice or mood.
iii)Redundancy and repetition: Unnecessary repetition of phrases or ideas.
iv)Lack of thematic depth: We found that there was surface-level treatment of complex subjects.
vi)Clichés and predictable language: Overused expressions and generic narrative patterns.

These issues were compared with examples of human-written literature on similar themes to highlight the contrasts.

3. Evaluation Criteria

We developed scored each piece on five parameters:
i)Emotional authenticity
ii)Narrative coherence
iii)Thematic richness
iv)Originality in expression
v)Structural flow

We rated each parameter on a scale of 1–5 and the average scores helped us to quantify the performance gaps between AI and human literary work.

4. Human Expert Review

A small panel of English literature students and faculty members participated in the process of evaluation. They reviewed the AI generated content and provided us the feedback on what felt unnatural, emotionally off and structurally weak. Their qualitative observations were instrumental in understanding the subtle ways in which AI misses the mark in creative writing and english literature.

5. Proposing Solutions

Based on the patterns of errors observed, the study showed many potential solutions. These were informed by:

i)Literature on AI emotional modeling and computational creativity.
ii)Feedback from our collegues.
iii)Current developments in AI-human collaboration tools.

6.We propose solutions like:

i)Integrating sentiment-awareness models to improve emotional alignment and emotional intelligence in AI models.
ii)Implementing human-in-the-loop editing systems for creative writing.
iii)Using contextual reinforcement learning to fine-tune literary style and depth over time.

7. Development of Emotional Intelligence in AI Models

i)This study also looks at ways on how we can improve AI's comprehension and reflection of human emotions in order to fix all the emotional flaws commonly seen in literature works produced by AI. The multi-step approach is:

ii)Emotion-Labeled Datasets: Use literary works that have been annotated for emotional tone, mood, and affective context in addition to grammar and structure to train AI. This teaches the model how emotions are organically incorporated into stories.

iii)Affective Computing Models: Apply concepts from the field of affective computing, which focuses on developing systems that can identify, understand, and replicate human emotions. These models can help AI match intended emotions with word choice, tone, and tempo.

iv)Contextual Sentiment Analysis: We should implement dynamic sentiment analysis that will adapt to the situations and progression in the story rather than applying static sentiment tags. This will allow AI to regulate emotional intensity based on narrative progression.

v)Reinforcement Learning with Human Feedback (RLHF): All the AI models should introduce a feature where human readers can provide real-time feedback on emotional accuracy and resonance of the literary works generated by them. Over the time, this feedback is going to help the AI improve its emotional abilities which will improve both emotional expression and reader connection of their works.

vi)Emotion-to-Expression Mapping: we should design internal frameworks that translate emotional states into literary techniques such as metaphor, pacing and imagery which will mirror how human writers express emotions creatively and emotionally.

By integrating these systems, AI can be trained not only to recognize but also to replicate the depth of human emotional expression in literary work.

8. Prompt Engineering for English Background people

Our research highlights the growing significance of prompt engineering which is the art of creating precise, creative and clear inputs that pushes AI tools to produce more meaningful and emotionally resonant literary outputs as a means of assisting English literature students and educators in making effective use of AI-generated content. A more thorough comprehension of tone, voice, narrative components and thematic intention is demanded by literature-based task prompts. A well-crafted prompts are to instruct an AI are "write a short story in the style of Virginia Woolf, focusing on internal monologue and emotional introspection" or "compose a free-verse poem about loss using winter as a metaphor and a melancholic tone. "By simply giving the AI more guidance, these prompts enable it to produce responses that are more favorable and aligning with human creativity and feels less robotic.

IV. RESULTS AND DISCUSSION

4.1 Many repetitive patterns of error that affect the depth, authenticity and emotional resonance of these works were observed after our analysis of AI-generated English literary content, which included essays, short stories and poems. Even though AI programs like GPT-4 and DeepSeek showed excellent grammar, vocabulary and sentence structures but they lack some essential components for creative writing that has an impact.

4.2 Emotional inconsistency is one of the most noticeable problems. The emotional tone of many AI-generated poems and stories is inconsistent throughout. For example, a poem intended to convey grief might suddenly change to a strangely upbeat tone or use metaphors that don't seem to fit the intended meaning.

4.3 We observed another common issue throughout our research and that was thematic shallowness. Human literature usually goes through themes with subtlety and layered meaning but on the other hand AI typically presents ideas in a flat and surface-level manner. It may touch on universal themes like love, loss, or hope, but it lacks the subtlety and undertone that come from lived experiences. These writings looks cliched or even robotic because they lack the originality and uniqueness that we, the humans have naturally.

4.4 We all can see the repeated phrasing and cliched language in these AI generated works. AI frequently reuses common expressions that results in texts that are generic and lack originality. The creative voice and originality that define influential literature are lost in this repetition, even though it can ensure readability.

4.5 All the errors and mistakes made by AI forces us to think about the current limitations of Artificial Intelligence. Taking the positive aspect into consideration, we can use these failures as an opportunity to improve and collaborate with AI to produce better works.

4.6 To address these issues, we propose the below solutions:

i) Improving Emotional Intelligence in AI: AI can be trained on datasets that are enriched with emotional annotations and literary tone labels. We can also Incorporate affective computing techniques and reinforcement learning with human feedback that is going to help AI models to improve consistent emotional flow in writing.

ii) Prompt Engineering to Provide Literary Depth: Giving AI well-structured, context-aware prompts can direct its output toward more emotionally compelling and significant results. Instead of asking an AI to "Write a sad poem," for instance, we could ask it to "Write a poem from the perspective of a grieving mother who finds hope in nature." This will give the story more emotional focus.

iii) Collaboration between humans and AI: AI should be viewed as a collaborator rather than a substitute. We maintain have emotional intelligence and creativity that machines currently lack but by letting authors and editors improve and polish AI drafts. We can use AI with it's full potential .

iv) Feedback Loops and Real-Time Learning: By incorporating real-time evaluations from human readers, AI can modify its word choice, tone, and structure in response to subjective feedback, transforming it from a generator into a learner.

v) Cultural Sensitivity and Context Modeling: If we start training AI on a variety of traditions, worldviews and opinions which will be helpful for it to avoid cultural biases and develop more natural language, especially in stories that deals with complex social themes and other sensitive topics. This is known as cultural sensitivity and context modeling(CSCM).

4.7 In conclusion, even though we all are aware about the potentials of Artificial Intelligence but at the same time we can't ignore the weakness it has in terms of emotions, tone, understanding and experiences which is very difficult to be adopted by AI. By identifying and resolving these issues through technological advancements and cooperative frameworks, we can transform AI into a tool that elevates rather than detracts from the craft of storytelling.

REFERENCES

[1] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. https://doi.org/10.1145/3442188.3445922

[2] Floridi, L., & Chiriatti, M. (2020). *GPT-3: Its Nature, Scope, Limits, and Consequences.* Minds and Machines, 30(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

[3] Kreps, S., McCain, R., & Brundage, M. (2022). *All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation.* Journal of Experimental Political Science, 9(1), 104–117. https://doi.org/10.1017/XPS.2021.28

[4] Liu, S., et al. (2022). *Designing Prompt Engineering Strategies for Large Language Models.* arXiv preprint arXiv:2212.09542.https://arxiv.org/abs/2212.09542

[5] Wang, Y., & Klinger, R. (2022). *Emotion-Infused Models for More Human-Like Text Generation.* In Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL).

[6] Roose, K. (2023). *The Uncanny Poetry of ChatGPT.* The New York Times. https://www.nytimes.com/2023/02/05/technology/chatgpt-poetry-writing.html

[7] OpenAI. (2023). *GPT-4 Technical Report.*https://openai.com/research/gpt-4

[8] Chakrabarty, T., Muresan, S., & Peng, N. (2020). *Generating Humorous Texts via Neural Networks.* In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). https://aclanthology.org/2020.emnlp-main.695/

[9] Yu, Z., et al. (2021). *How Emotions are Made: Building Emotionally Aware AI.* ACM Transactions on Human-Robot Interaction (THRI), 10(3).https://doi.org/10.1145/3460320

[10] Veale, T. (2020). *Creativity and Artificial Intelligence: A Short Introduction.* In The Routledge Companion to Creativity.