**Seminar Project Report**

**Title: - Elara: Mood-Based Story and Image Generator using LLM and Stable Diffusion.**

1. **Introduction:**

In the era of Generative AI, merging textual creativity with visual generation opens up powerful possibilities. This project, titled **"Elara"**, integrates two generative AI models—**a Large Language Model (LLM)** and **a text-to-image diffusion model**—to deliver a unique experience where a user enters a prompt, chooses a story mood (e.g., fantasy, sad, funny), and the system generates both:

* A short story that aligns with the mood
* A visual image based on the prompt

The purpose is to demonstrate how **multimodal AI systems** can create immersive storytelling experiences by combining **natural language generation** with **image synthesis**.

1. **Tools and Technologies Used:**

| **Component** | **Technology Used** |
| --- | --- |
| Language Model | LLaMA3 70B via Groq API |
| Image Generation | Stable Diffusion v1.5 (Hugging Face Diffusers) |
| UI/Frontend | Gradio (Python-based UI framework) |
| Platform | Google Colab (for prototyping and testing) |
| API Key Management | google.colab.userdata (Colab secrets) |
| Programming Language | Python 3.11 |

1. **Methodology**

**System Design Overview**

The system is composed of two AI models:

1. **LLM (LLaMA3)**: Generates a 4–6 line short story from the user’s prompt and selected mood.
2. **Stable Diffusion**: Creates a visual image from the same prompt using latent diffusion.

**Workflow:**

* **Step 1**: User provides a text prompt (e.g., "a lonely robot in the desert")
* **Step 2**: User selects a mood: Fantasy, Funny, Sad, Sci-Fi, Adventure, or Romantic
* **Step 3**: The chatbot (Elara) responds with a story matching the mood using Groq API
* **Step 4**: At the same time, the prompt is passed to the Stable Diffusion model to generate an image
* **Step 5**: Both outputs are shown side by side using Gradio UI

1. **Introduction to System Functions:**

The system has the following core functions:

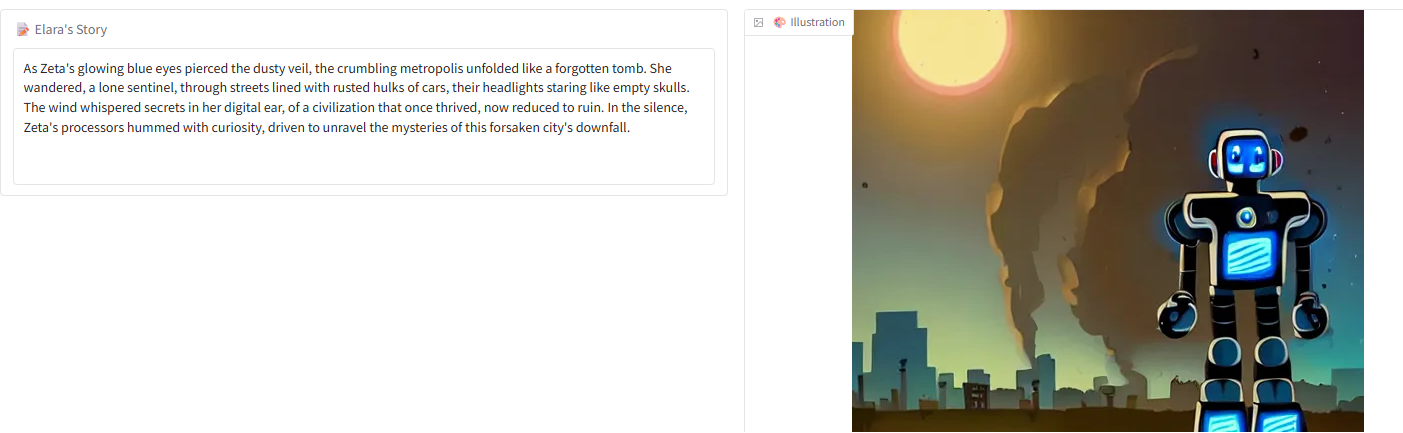
| **Function** | **Description** |
| --- | --- |
| **Prompt Input** | User enters a short creative idea (e.g., “a cat flying a spaceship”) |
| **Mood Selection** | User chooses a story mood: Fantasy, Funny, Sad, Sci-Fi, Adventure, or Romantic |
| **Story Generation** | LLaMA3 generates a 4–6 line short story based on prompt and mood |
| **Image Generation** | Stable Diffusion v1.5 generates a visual image from the same prompt |
| **UI Display** | Outputs are shown side-by-side on a web app built with Gradio |

1. **Preliminary Approach to the System:**

The system is developed using the following steps:

1. **LLM Integration**:
   * Used Groq API to call LLaMA3 (70B) model
   * Defined the system message for storytelling mode
   * Used user prompt + mood to generate a custom story
2. **Image Generator Setup**:
   * Used Hugging Face Diffusers to load Stable Diffusion v1.5
   * Generated image from user prompt
   * Adjusted generation using guidance scale and inference steps
3. **Frontend Integration**:
   * Built a unified interface using Gradio Blocks
   * Included dropdown for mood, sliders for image config
   * Outputs displayed in real-time
4. **Testing and Optimization**:
   * Multiple prompts tested for quality
   * System debugged for runtime delays and GPU issues
5. **Results:**

Example [Prompt: - **A robot exploring an abandoned city**].

Mood: Sci-Fi. **Output Story: -**

**Conclusion:**

This project successfully demonstrates how **two distinct Gen AI tools**—LLM and Diffusion—can be fused to create a creative assistant that not only tells stories but also brings them to life visually. It showcases the versatility and accessibility of multimodal Gen AI and encourages future exploration in **interactive storytelling**, **AI companions**, and **personalized learning/entertainment tools**.

1. **References**

* Groq API: <https://console.groq.com/>
* Hugging Face Diffusers: [https://huggingface.co/docs/diffusers](%20https:/huggingface.co/docs/diffusers)
* Stable Diffusion v1.5: <https://huggingface.co/runwayml/stable-diffusion-v1-5>
* Gradio: <https://www.gradio.app/>

🔗 **MY GitHub Repository:** <https://github.com/AnwarHussain-1011/elara-moodgen-elara-ai-story-art->