**PROJECT TOPIC : Text-To-Image Generation**

**Specialization: CSE**

**Project Group Members:**

1. Yash Patel (D-52 / 2115001156.)

**Project Mentor:** Dr. Sabita Arora

**Objective:** The objective of this project is to develop a web-based AI-powered text-to-image generation system that allows users to convert natural language prompts into high-quality images. By leveraging deep learning models, particularly the Stable Diffusion model, the platform enables creative visualizations that reflect the semantic meaning of textual inputs. This tool is designed for users such as digital artists, designers, educators, and researchers who seek to generate visual content without needing design expertise. The seamless integration of the frontend and backend ensures an interactive and responsive experience, providing real-time image generation with a focus on accuracy, creativity, and user control.

**Tools required:**

**Hardware Requirements:**

* A development PC or laptop with a dedicated NVIDIA GPU (recommended: RTX 3060 or above)
* Stable internet connection
* Monitor with good display resolution for visual evaluation

**Software Requirements:**

* **Frontend:**
  + React.js (for building the user interface)
  + HTML, CSS, JavaScript
  + Axios (for API integration)
* **Backend:**
  + Node.js and Express.js (for routing and handling requests)
  + Python 3.x (for AI model execution)
  + Torch (PyTorch library for model loading)
  + diffusers and transformers (from Hugging Face for Stable Diffusion)
  + Mongoose and MongoDB (for storing user prompts and metadata)
* **Other Tools:**
  + Git & GitHub (for version control)
  + Postman (for API testing)
  + VS Code (as the IDE)

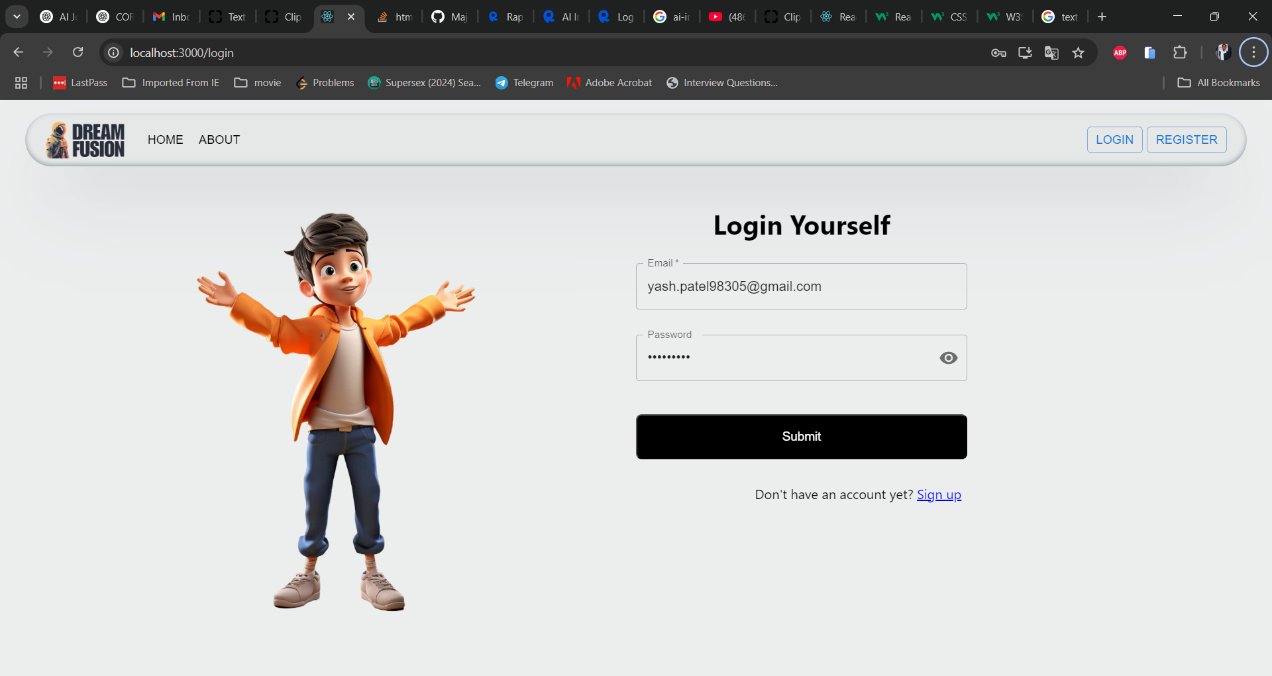
**Abstract**

This project, titled **“Text-to-Image Generator”**, is an AI-powered web application designed to convert textual descriptions into realistic images using deep learning techniques. Built on the MERN stack (MongoDB, Express.js, React.js, Node.js) and integrated with Python-based Stable Diffusion models, the system demonstrates the practical utility of transformer-based diffusion models in bridging the gap between language and vision. Users can enter descriptive prompts through a clean and intuitive interface, triggering the backend to process and forward the input to a pretrained model that generates a semantically aligned image.

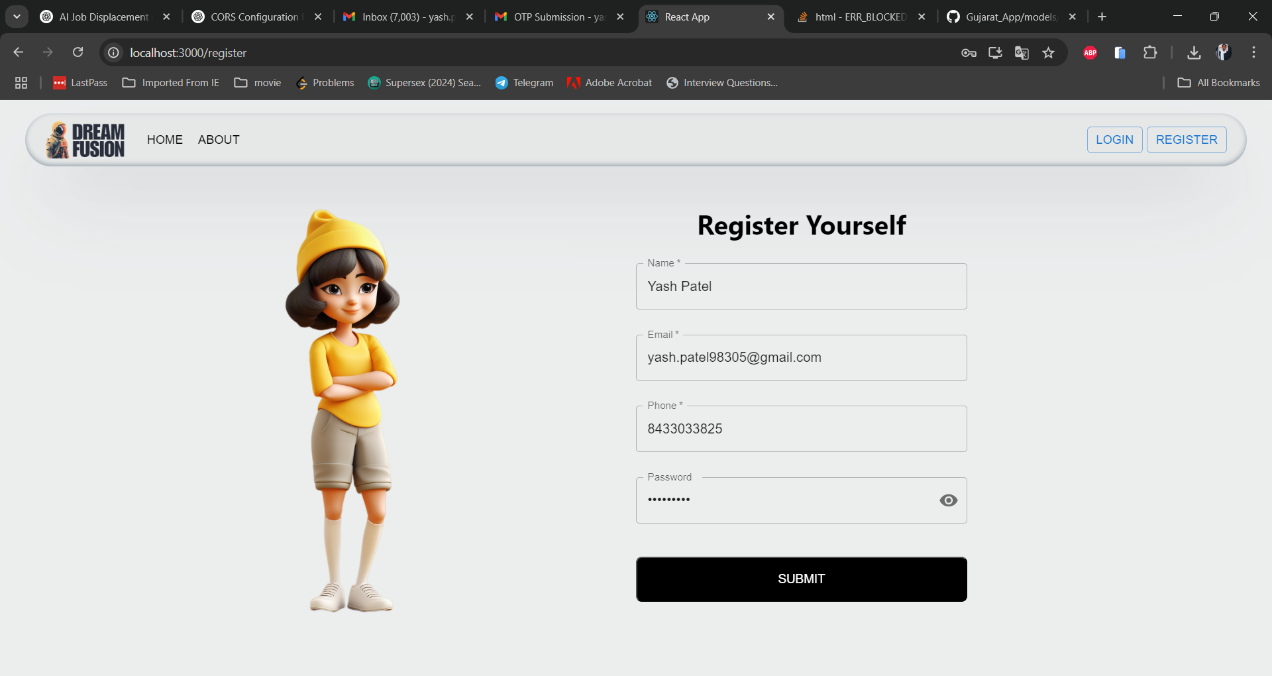
The application serves a variety of use cases—ranging from content creation and education to marketing and storytelling—by allowing dynamic image synthesis with minimal user effort. With key features like GPU acceleration, prompt filtering, and customizable output sizes, the system emphasizes both functionality and performance. This AI-driven platform not only showcases the power of generative models but also opens avenues for future enhancements like style customization, fine-tuned model deployment, and multilingual support.

**Outcomes**

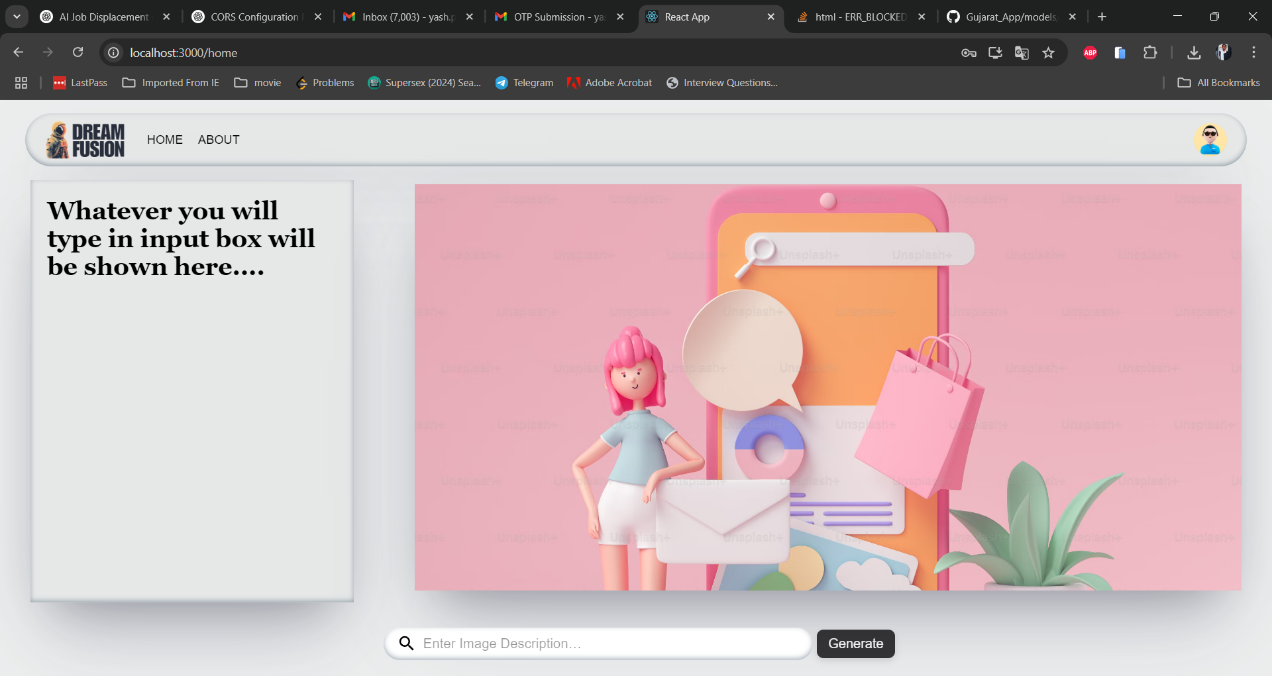
**Login Page**



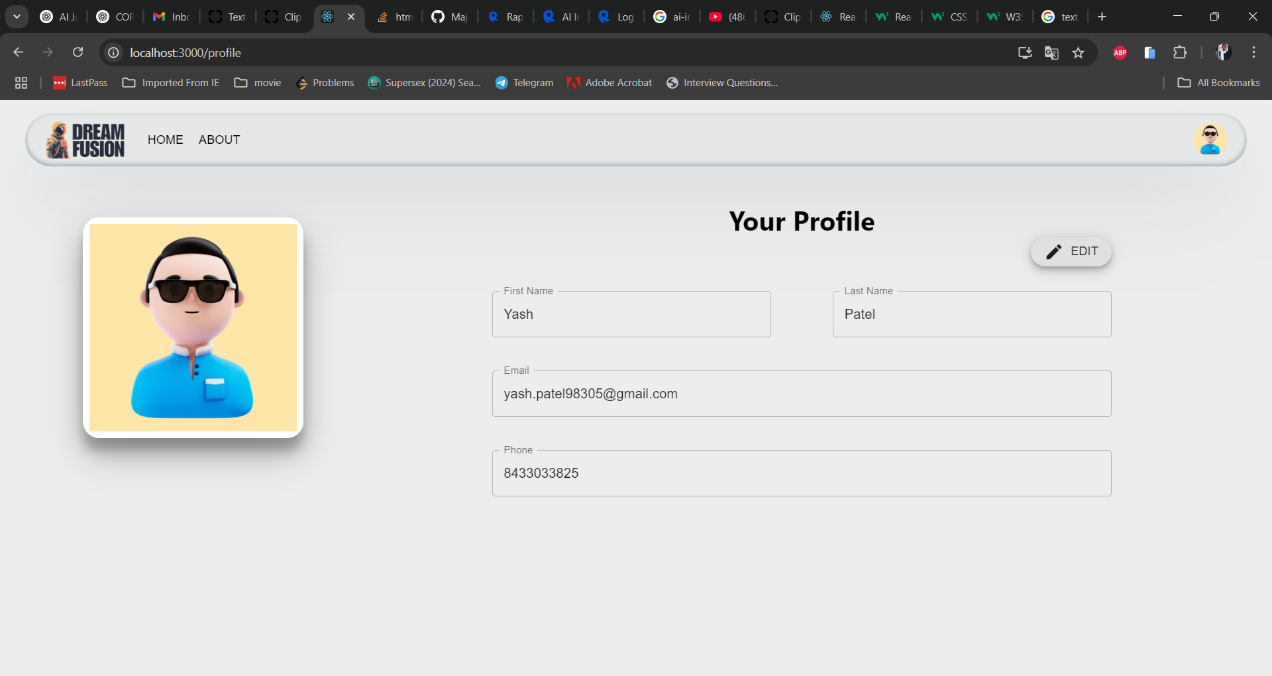
**Register Page**



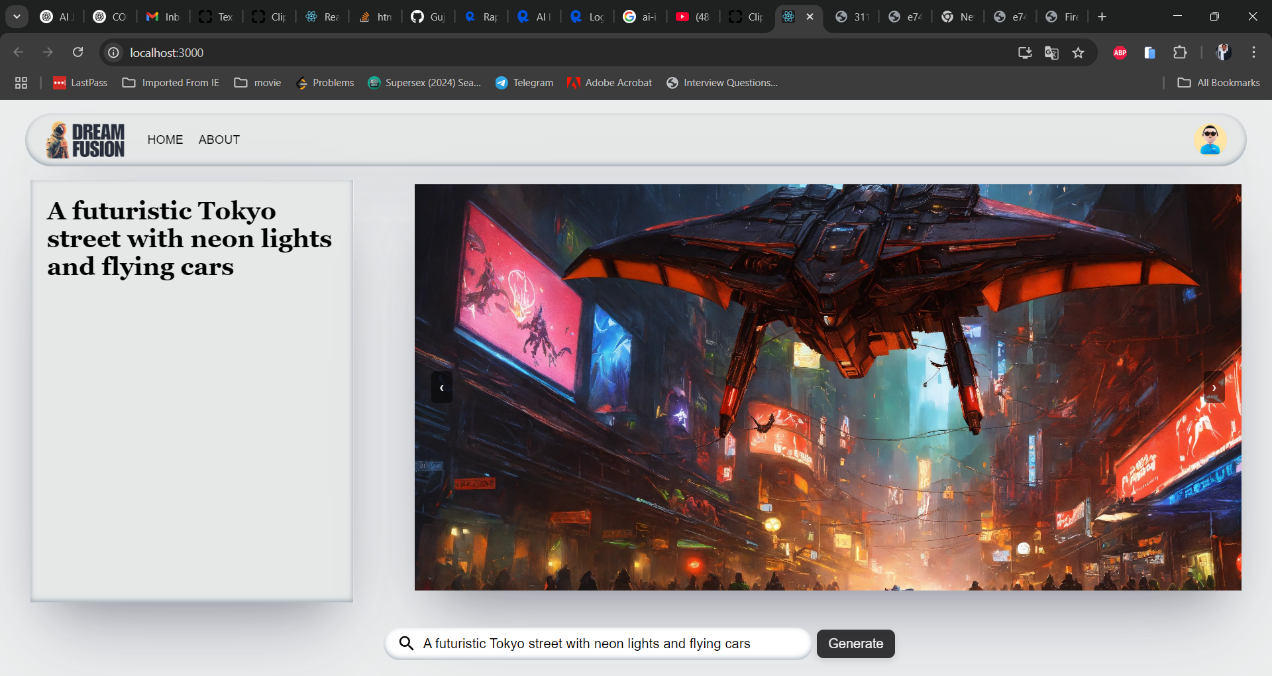
**Home Page**



**Profile Page**



**Prompt and Image**



**History Management**

