PAPER NAME

Capstone Project P46 Report (1).docx

WORD COUNT

12340 Words

CHARACTER COUNT

75243 Characters

PAGE COUNT

78 Pages

FILE SIZE

12.2MB

SUBMISSION DATE

May 9, 2024 4:33 PM GMT+5:30

REPORT DATE

May 9, 2024 4:34 PM GMT+5:30

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 3% Internet database
- 1% Publications database
- Crossref database
- Crossref Posted Content database
- 6% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Small Matches (Less then 8 words)

Project Report

on

AI-Based Therapeutic Image Generation Using Stable Diffusion

Submitted by

Project Members

Yash Honrao 1032200191

Prachiti Chitta 1032200527

Viswas Haridas 1032200739

Vrishani Shah 10322001102

Under the Internal Guidance of

Prof. Dr. Shamla Mantri

School of Computer Engineering and Technology

MIT World Peace University, Kothrud,

Pune 411 038, Maharashtra - India

2023-2024

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

DEPARTMENT OF COMPUTER ENGINEERING AND TECHNOLOGY

C E R T I F I C A T E

This is to certify that,

Yash Honrao

Prachiti Chitta

Viswas Haridas

Vrishani Shah

of BTech. (Computer Science & Engineering) have completed their project titled *"AI-Based Therapeutic Image Generation Using Stable Diffusion"* And have submitted this Capstone Project Report towards the fulfillment of the requirement for the Degree-Bachelor of Computer Science & Engineering (BTech-CSE) for the academic year 2023-2024.

[Dr. Shamla Mantri]
Project Guide
School of CET
MIT World Peace University, Pune

[Dr. Balaji Patil]
Program Director
Department of CET
MIT World Peace University, Pune

Internal Examiner:_____

External Examiner:_____

Date: 09/05/2024

Acknowledgement

We extend our deepest gratitude to Dr. Shamla Mantri for her invaluable guidance, mentorship, and unwavering support throughout the journey of the "AI-Based Therapeutic Image Generation Using Stable Diffusion" project. Her profound expertise in the field, combined with her patient and insightful approach, immensely contributed to the depth and success of this endeavor. Her ability to encourage critical thinking and foster an environment of academic rigor has not only aided in the successful completion of this project but also significantly enriched my professional and personal growth.

The support and encouragement from the CET department at MIT World Peace University, Pune, coupled with the collaborative spirit and assistance from colleagues and peers, have been instrumental in overcoming challenges and achieving the project goals. This journey has been a testament to the power of teamwork, mentorship, and academic pursuit; we are eternally grateful for this.

Lastly, we thank our parents for their years of unyielding love and encouragement. They always wanted the best for us, and we admire their sacrifice and determination.

Yash Honrao

Prachiti Chitta

Viswas Haridas

Vrishani Shah

Abstract

In the contemporary landscape of mental health, the urgency for accessible and personalized interventions is heightened by the increasing prevalence of anxiety disorders. This project introduces a pioneering artificial intelligence (AI) system designed to generate personalized therapeutic visuals and narratives, catering to individual preferences and psychological needs. The system leverages cutting-edge technologies including reinforcement learning (RL), generative adversarial networks (GANs), and image generation algorithms such as Stable Diffusion. By integrating user inputs—preferences for specific themes such as animals, shapes, and nature—the AI tailors its output to enhance comfort and emotional support. This approach is underpinned by principles of positive psychology, mindfulness, and cognitive-behavioral therapy, further enriched by the integration of thematic music and AI-generated storytelling. The system's unique feature, an SOS functionality, allows users to instantly generate therapeutic content during acute anxiety episodes, providing immediate relief. This review synthesizes various methodologies from existing literature to propose a model that not only adapts to user preferences over time but also contributes significantly to the field of mental health by offering a non-invasive, adaptive, and user-centered tool for anxiety management.

Keywords: artificial intelligence, mental health, anxiety management, personalized therapy, image generation, reinforcement learning, generative adversarial networks, user-centered design.

List of Figures

List of Tables

Contents

Chapter 1
Introduction

This project, 'AI Based Therapeutic Image Generation Using Stable Diffusion,' delves into the development of a novel system for generating therapeutic images using stable diffusion techniques. By harnessing the power of artificial intelligence, this endeavor aims to produce visually soothing and emotionally calming images. Through the utilization of stable diffusion algorithms, the project endeavors to create a platform capable of generating images that promote relaxation, mindfulness, and emotional well-being. The ultimate goal is to explore the potential of AI in providing therapeutic support through the creation of visually compelling and emotionally resonant imagery, thereby contributing to the advancement of mental health technology."

1.1 Problem Statement:

Develop an AI system that creates personalized therapeutic audio-visuals for individuals facing challenges such as anxiety. The system could analyze user input, including their preferences (eg. animals, shapes, etc.), and generate evolving images designed to provide comfort, support, and coping strategies. The images might incorporate elements of positive psychology, mindfulness, and cognitive-behavioral therapy along with relevant music. Users can engage with AI-generated stories as part of a therapeutic process, creating a unique and tailored experience that aids mental well-being. This project combines generative AI with a serious societal impact, offering a potential tool for mental health support and intervention.

1.2 Problem Introduction:

Being happy in life and achieving goals is needed the most in our lives. However mental health issues like anxiety, stress, and depression cause a hindrance in achieving this goal. Emotions that are out of control and come abruptly create problems and chaos in one’s life. Medical conditions like continuous anger and impatience can cause heart problems and in extreme cases cause even the bursting of an artery. Managing these emotions in an emergency situation and controlling them immediately can help in resolving such problems.

With an increase in digital solutions, using Artificial Intelligence(AI) in the prediction, detection, and treatment has become imperative. Smartphone applications now have an inclusion of AI to aid in helping patients. Gathering the usage data of users helps in data mining and giving inputs on how AI can help in solutions. Even language is now used to harness the best ways to understand mental health using Natural Language Processing (NLP). The major methods in which AI has aided have been personal sensing or digital phenotyping, natural language processing of clinical content, and social media content and chatbots. Digital phenotyping and personal sensing collect data from devices which can then be utilized in machine learning to understand and predict psychological outcomes and mental health conditions. A person's method of talking helps to understand if someone is suffering from any condition. Chatbots use rule-based techniques and matching keywords to diagnose mental conditions.

Guided imagery or visualization therapy helps individuals imagine a scene and how they would be a part of it including their reactions to the images. The images help people to control how they react, change their responses to the situations, and act as a coping mechanism. It provides multiple ways to manage stress and anxiety.

In this project, we propose a combination of Artificial Intelligence and Stable Diffusion.

1.3 Project Aim:

The aim of the project "AI-Based Therapeutic Image Generation Using Stable Diffusion" is to develop an innovative platform that leverages stable diffusion algorithms and artificial intelligence techniques to create a diverse range of therapeutic images. These images are intended to foster relaxation, reduce stress, and promote emotional well-being. By harnessing the power of AI, the project seeks to generate visually compelling and emotionally resonant imagery that can be tailored to individual preferences and needs. Through this endeavor, the aim is to provide a versatile tool for mental health professionals, individuals seeking self-care, and those undergoing therapy, offering them access to a rich repository of calming and soothing visual content designed to support their well-being journey.

Chapter 2

Literature Survey

2.1 Literature Review of Existing Work:

Luxton's (2015) in Artificial Intelligence in Behavioral and Mental Healthcare: Review in AI services can help with mental health treatment and related applications. Sometimes even the number of people practicing healthcare can be insufficient to the people suffering from it and thus they are unable to receive help. This resource gap can be addressed using virtual techniques, especially using AI. The interaction with AI gives it a more human aspect and it also helps in making the process more helpful instead of just videos or websites. Even geographically, in remote areas, it helps to get help where no practitioners may be available.

Inner pictures have a very strong effect on people and can influence them greatly. They have made athletes win races and cure inner problems like headaches and tension. To achieve big goals, a combination of physical and mental training is an important factor. The training of the mind to remove negative images and imagine victory images is important. The internal images are done by asking the patient to cover a gap that is formed when the computer shows two images, usually when someone is suffering from depression, they get negative and destructive images to close the gap. Even in company settings, when employees were made to sit with therapists, they imagined the company problems in their minds and could not let go of them until the solution was found.

Rusch, Grunert, Mendelsohn, and Smucker's (2000) review in Cognitive and Behavioral Practice: Images can play an important role in emotional distress. Intrusive images can cause emotional distress. Images can hold certain memories and act as triggers. Using distraction techniques, therapy interventions, and prolonged exposure can cure these problems. Image rescripting can also be done where the therapist guides the person to manipulate or modify these images in their mind. Patients visualize the anxiety-provoking scene and how they would react positively to it.

Byrnes' (1996) review in The Effect of Audio, Video and Paired Audio-Video Stimuli: Audio and visual aids have been used commonly to ease anxiety and stress. Studies on college students have shown reduced anxiety using music and imagery techniques. Even patients who would be undergoing surgery seemed to become calmer when given music and image therapy.

There are many tapes available termed as “anxiety-reducing” or “relaxation” which are also sold to hospitals and healthcare facilities. Results of studies show that using audio-visuals reduces stress levels and helps calm a person.

Hamdoun, Monteleone, Bookman, and Michael’s review in AI-Based and Digital Mental Health Apps: Balancing Need and Risk: After the COVID-19 pandemic, the cases of mental health problems have increased dramatically and multiple gaps in mental health services have been found. Due to this the need for more mental health applications has come up, especially chatbots.

Clark’s (2001) review for SOS Help for Emotions: Managing Anxiety, Anger, and Depression: In "SOS Help for Emotions: Managing Anxiety, Anger, and Depression," Lynn Clark offers a practical guide aimed at helping readers develop effective strategies to manage complex emotional challenges. The book, through its illustrated and easily accessible format, introduces cognitive-behavioral techniques that empower individuals to alter their emotional responses by changing their thought patterns. The guide is particularly noted for its straightforward language and actionable advice, making psychological strategies accessible to a broader audience without the need for professional intervention. This work is beneficial for those seeking to enhance their emotional intelligence and gain control over anxiety, anger, and depression, promoting better mental health and well-being.

Simon D’Alfonso’s (2020) review for AI in mental health, Current Opinion in Psychology: Simon D'Alfonso's "AI in Mental Health" delves into the emerging role of artificial intelligence within the realm of psychological health services. This 2020 article in Current Opinion in Psychology critically examines the potential and challenges of AI applications in mental health diagnosis and treatment. D’Alfonso discusses AI-driven predictive models that enhance the accuracy of mental health assessments and the development of personalized treatment plans. The paper also addresses ethical considerations, such as privacy and the need for transparency in AI applications. This work provides a crucial insight into how AI tools can complement traditional psychological approaches, potentially leading to more efficient and accessible mental health care.

S. Zhou, J. Zhao, and L. Zhang’s (2022) review for Application of Artificial Intelligence on Psychological Interventions and Diagnosis: This article provides a comprehensive overview of the intersection between artificial intelligence (AI) and psychology, focusing on its application in psychological interventions and diagnosis. Published in Front. Psychiatry, the paper evaluates the current state of AI tools that aid in diagnosing mental health disorders and delivering therapeutic interventions. It highlights the effectiveness of AI in extending the reach

of psychological services, especially in under-resourced areas, and enhancing the precision of psychological assessments. The authors also consider the ethical implications of AI, including privacy concerns and the risk of dehumanization. This review underscores the transformative potential of AI in reshaping mental health practices.

D. B. Dwyer, P. Falkai, and N. Koutsouleris' (2018) review for Machine Learning Approaches for Clinical Psychology and Psychiatry: In their 2018 publication in the Annual Review of Clinical Psychology, Dwyer, Falkai, and Koutsouleris explore the application of machine learning (ML) techniques in the field of clinical psychology and psychiatry. The paper discusses how ML models are revolutionizing the detection, diagnosis, and treatment of psychiatric disorders by analyzing complex datasets that traditional methods cannot. It provides an in-depth analysis of the potential for these technologies to predict treatment outcomes, personalize therapy approaches, and uncover new insights into the biological underpinnings of mental disorders. The review also critically examines the limitations and ethical considerations of ML in clinical settings, providing a balanced view of its benefits and challenges.

D'Alfonso's (2017) review for Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health: This 2017 article by D'Alfonso et al., published in Front. Psychol. introduces an innovative approach to youth mental health therapy through the use of Artificial Intelligence-assisted online social platforms. The study presents the design and efficacy of AI tools that facilitate social therapy for young individuals, enhancing engagement and providing real-time therapeutic interventions. By integrating AI with therapeutic practices, the study demonstrates improvements in accessibility and customization of mental health care for youth. The paper significantly contributes to the understanding of how AI can be harnessed to support mental health in a demographic prone to technology engagement, offering promising directions for future research and implementation.

Mataghare, Aote, and Hablani's (2023) review for Text to Image Generation using Stable Diffusion: This paper explores the capabilities of diffusion models in the domain of image generation, particularly focusing on text-to-image synthesis using Stable Diffusion. The authors highlight the efficiency of latent diffusion models (LDMs) which operate in a compressed latent space rather than direct pixel manipulation. This approach significantly reduces the computational demands and speeds up the image generation process without sacrificing the quality of the generated images. The paper also discusses the integration of cross-attention layers within the diffusion model framework, which enhances the model's ability to handle diverse and complex conditioning inputs, such as textual descriptions and bounding boxes. This advancement pushes the boundaries of image synthesis, presenting a

scalable solution for generating high-fidelity images from textual descriptions.

Zhou, Zhu, and Naka's (2024) review for Text Image Generation In DCGAN and Stable Diffusion Model: This research delves into comparing the efficacy of DCGAN and Stable Diffusion models for text-to-image applications. The study utilized the CLIP ViT-B/32 model for text embeddings and trained the models using the MNIST dataset. The findings indicate that while DCGAN models generally produce higher-quality images, Stable Diffusion models excel in terms of computational efficiency. The authors suggest potential enhancements to the diffusion models by integrating aspects of the DCGAN discriminator, aiming to improve the denoising capabilities of the diffusion models. This paper contributes to understanding the trade-offs and synergistic potentials between generative adversarial networks and diffusion models in text-to-image synthesis.

Review for StableDiffusionIntelliSys2023: The authors introduce innovative enhancements to the Stable Diffusion model by integrating Hypernetworks and DreamBooth. This integration aims to improve the model's adaptability and efficiency in generating images from minimal training data while ensuring diversity and inclusion in the generated images. The paper highlights the practical applications of these advancements in various domains such as smart cities, indicating significant improvements in personalization and context-aware image generation. The authors also discuss the broader impacts of these technologies, emphasizing their potential to enhance AI-driven applications across different sectors.

Hidalgo, Salah, Jetty, Jetty, and Varde's (2023) review for Personalizing Text-to-Image Diffusion Models: This paper proposes a novel approach to personalize text-to-image generation using Stable Diffusion by employing Hypernetworks and DreamBooth. The study focuses on the adaptability of these models to incorporate a broad range of real-world data and generate diverse, high-quality images. It addresses the challenges of pre-trained models' limitations by allowing for the introduction of new images into the model with minimal additional training. This approach significantly enhances the application potential of text-to-image models in areas requiring high levels of customization and precision.

Review for Exploring Text-to-Image Diffusion Models (2023): The document "Exploring Text-to-Image Diffusion Models" investigates the integration of advanced diffusion models for text-to-image conversion, emphasizing the enhancement of model responsiveness to diverse textual inputs. It explores the application of these models in various AI-driven tasks, highlighting the improvement in generating contextually relevant and visually appealing images from text descriptions. The paper serves as a crucial step towards realizing the full potential of diffusion models in practical applications, providing a foundation for future

innovations in AI and image processing technologies.

2.2 Limitations & Research Gaps of Existing Work:

The project addresses the critical shortcomings in existing AI therapeutic systems, emphasizing the necessity for dynamic content adaptation to enhance user experiences. Current systems often lack the ability to adjust content dynamically based on real-time emotional states and evolving preferences, leading to a significant gap in meeting the diverse needs of individuals. To bridge this gap, there is a pressing need for the development of AI systems capable of continually updating and modifying therapeutic content in response to user feedback and changing emotional needs. Additionally, the report underscores the importance of personalization in therapy interventions, highlighting the limitations of current technologies in incorporating extensive individual user inputs to create truly tailored therapeutic experiences. By leveraging sophisticated AI models that can understand and integrate complex user data points such as preferences, behaviors, and emotional feedback, it becomes possible to deliver personalized interventions that resonate deeply with users. Furthermore, the report emphasizes the necessity of incorporating mental health dynamics into AI systems, as existing static content fails to address the fluid and variable nature of mental health conditions. By recognizing and adapting to the various coping strategies required by different individuals, AI systems can offer more effective mental health support and intervention, ultimately improving overall well-being.

Chapter 3
Problem Statement

3.1 Project Scope:

The project "AI Based Therapeutic Image Generation Using Stable Diffusion" entails the creation of an innovative platform that leverages stable diffusion algorithms and artificial intelligence techniques to generate therapeutic images. The scope encompasses several key components, including the development of robust algorithms for stable diffusion image generation, the design of a user-friendly interface for seamless interaction, and the implementation of deep learning models to enhance image quality and emotional resonance. Additionally, the project involves the collection and analysis of user feedback to refine and personalize the generated images according to individual preferences and emotional needs. Evaluation metrics will be established to assess the effectiveness of the generated images in promoting relaxation, reducing stress, and enhancing overall well-being. Through comprehensive testing and iteration, the project aims to deliver a cutting-edge solution for therapeutic image generation that offers personalized support and fosters emotional wellness.

3.2 Project Assumptions :

Firstly, it is assumed that stable diffusion algorithms will proficiently generate visually soothing and emotionally resonant images, drawing upon previous research affirming their efficacy in this domain. Additionally, the assumption is made that user feedback will furnish invaluable insights for refining and personalizing the generated images, thereby iteratively enhancing user satisfaction and emotional resonance. Furthermore, it is assumed that leveraging deep learning models will amplify the quality and diversity of the generated images, surpassing traditional methods in realism and emotive impact. A user-friendly interface design is anticipated to facilitate intuitive interaction and customization of image generation parameters, fostering enhanced user engagement and empowerment. Lastly, it is assumed that meticulously defined evaluation metrics will accurately gauge the effectiveness of the generated images in fostering relaxation and reducing stress, providing a robust basis for iterative improvement and validation. These assumptions serve as foundational pillars guiding

the project's trajectory, shaping its planning, execution, and evaluation processes.

3.3 Project Limitations:

While the "AI Based Therapeutic Image Generation Using Stable Diffusion" project holds promise for advancing mental health support through innovative technology, it also faces several limitations that warrant consideration:

1. Algorithmic Limitations: The effectiveness of stable diffusion algorithms and generative models like GANs in generating therapeutic images may vary depending on factors such as dataset quality, model architecture, and hyperparameter tuning. Limited access to high-quality training data or computational resources could constrain the performance and diversity of generated images.

2. Subjectivity of Therapeutic Response: The subjective nature of therapeutic responses and preferences poses a challenge in creating universally effective and emotionally resonant images. Individual differences in emotional states, cultural backgrounds, and personal experiences may influence the perceived effectiveness of generated images, leading to variability in user satisfaction and outcomes.

3. Ethical Considerations: The use of AI-generated images for therapeutic purposes raises ethical considerations regarding user privacy, consent, and potential misuse of personal data. Ensuring transparency, informed consent, and responsible data handling practices is essential to maintain user trust and ethical integrity throughout the project lifecycle.

4. User Accessibility: Limited accessibility to technology or digital platforms among certain user groups, such as older adults, individuals with disabilities, or those in low-resource settings, could hinder the reach and impact of the therapeutic image generation system. Addressing accessibility barriers and ensuring inclusivity is crucial to maximize the system's potential benefits for all users.

5. Validation and Clinical Efficacy: While the project aims to leverage AI for therapeutic purposes, rigorous validation and clinical efficacy studies are needed to establish the effectiveness and safety of AI-generated images as a complementary mental health

intervention. Collaborating with mental health professionals and conducting empirical research is essential to validate the therapeutic benefits and inform evidence-based practices.

6. Maintenance and Sustainability: The long-term maintenance and sustainability of the project, including software updates, server maintenance, and user support, pose ongoing challenges. Adequate resources, funding, and organizational commitment are necessary to ensure the continued availability and relevance of the therapeutic image generation system beyond the initial development phase.

3.4 Project Objectives :

The primary objective of this project is to harness the synergies between psychology and artificial intelligence (AI) to develop an innovative platform for therapeutic image generation. Through the integration of psychological principles and AI techniques, the project aims to create a repository of visually compelling and emotionally resonant images that promote relaxation, reduce stress, and enhance overall well-being. By leveraging AI algorithms and deep learning models, the objective is to generate images that resonate with individuals on a deeply personal level, addressing their unique emotional states, preferences, and needs. Additionally, the project seeks to advance our understanding of the psychological mechanisms underlying therapeutic imagery and its impact on mental health by incorporating insights from psychology into the design and evaluation of the image generation process. Ultimately, the project aims to pioneer a novel approach to mental health support, leveraging cutting-edge technology and psychological expertise to empower individuals on their journey towards emotional wellness.

Chapter 4
Project Requirements

4.1 Resources :

4.1.1 Software Requirements

Algorithms and Techniques: The project employs a combination of reinforcement learning algorithms and generative models (GANs, Stable Diffusion) to develop a user recommendation system and an image generation system. Reinforcement learning is used to optimize the system's responses based on user interactions and feedback. Generative models are utilized to create personalized therapeutic visuals that adapt to user preferences.

4.1.2 Software Development Environment

1. Programming Language: Python is used for both model development and backend due to its simplicity and extensive libraries. It's efficient for data manipulation and integrates well with deep learning frameworks.
2. Deep Learning Frameworks: TensorFlow or PyTorch are chosen for their robustness in implementing complex models like Stable Diffusion or GANs, offering powerful tools for training and deploying neural networks.
3. Backend Framework: Python Flask on Google Colab provides a lightweight yet scalable server environment. This setup ensures seamless communication with the mobile app through a REST API, leveraging Google's cloud infrastructure.
4. Database: Firebase Firestore is employed as a flexible NoSQL database. It offers real-time data syncing, scalability, and ease of integration with mobile and web applications.
5. Frontend Framework: Flutter with Material UI is ideal for developing visually appealing and responsive mobile interfaces. Its single codebase for iOS and Android reduces development time and effort.
6. Authentication: Firebase OAuth is used for user authentication, providing a secure and straightforward method to manage user identities across platforms.
7. Version Control: Git is utilized for managing the codebase and facilitating collaboration among developers. It ensures version control, code review, and tracking of changes

throughout the project's lifecycle.

4.1.3 Software Deployment and Hosting

For software deployment and hosting, the project utilizes Ngrok for temporary tunneling of the local Flask server to a publicly accessible URL during the development and testing phases. Ngrok facilitates easy sharing and testing of the application with stakeholders or collaborators without the need to deploy it on a public server. Additionally, the project leverages Google Collab to host the Flask server in a cloud-based environment. By deploying the server on Google Collab, the application benefits from scalability, reliability, and accessibility, enabling seamless access to the therapeutic image generation system from any device connected to the internet. These deployment strategies ensure efficient development, testing, and distribution of the application while maintaining accessibility and ease of use for end-users downloading the app from the Google Play Store.

4.1.4 Hardware Requirements:

1. Computer or server with sufficient processing power and memory to train and run the machine learning models.
2. Graphics Processing Unit (GPU) with CUDA support for accelerated training of deep learning models like Stable Diffusion or GAN.
3. This breakdown ensures that the content is organized logically under the relevant subheadings, providing clarity and coherence to the document.

4.2 Requirements Rationale :

S. No.	Requirement	Rationale
1	Programming Language	Python is chosen as the development language due to its versatility, extensive libraries for machine learning, and ease of use. It provides a robust foundation for model development and backend implementation.
2	Deep Learning Frameworks (TensorFlow and PyTorch)	TensorFlow and PyTorch are leading deep learning frameworks known for their flexibility, performance, and extensive community support. They offer comprehensive tools and resources for implementing complex deep learning models such as Stable Diffusion or GAN.
3	Backend Framework (Python Flask server on Google Collab)	Python Flask is selected as the backend framework for its lightweight and flexible nature, making it suitable for hosting a REST API. Google Collab provides a cloud-based environment for hosting the Flask server, ensuring scalability and accessibility without the need for a dedicated infrastructure setup.
4	Database (Firebase Firestore)	Firebase Firestore is chosen as the database for its real-time database capabilities, scalability, and ease of integration with other Firebase services. It allows efficient storage and retrieval of user data, including preferences and feedback.
5	Frontend Framework (Flutter, Material UI)	Flutter is selected for developing the frontend due to its cross-platform compatibility, fast development cycles, and rich set of customizable UI components. Material UI provides a cohesive design language and UI elements for a visually appealing and intuitive user experience.
6	Authentication (Firebase OAuth)	Firebase OAuth is used for authentication to provide secure access to the application and its features. It offers robust authentication mechanisms, including

		email/password, social login, and multi-factor authentication, ensuring user data privacy and security.
7	Version Control (Git)	Git is utilized for version control to manage the codebase, track changes, and facilitate collaboration among team members. It enables seamless code sharing, branching, and merging, ensuring code quality and project stability throughout the development lifecycle.

Table 4.2.1: Requirements Rationale

4.3 Risk Management :

S. No.	Risk	Mitigation	Risk Intensity
1.	Technical Risks: Technical challenges in implementing deep learning models (e.g., Stable Diffusion or GAN) could lead to delays in project timelines.	Conduct thorough research and prototyping to understand the complexities involved. Collaborate with experts or seek assistance from online communities for troubleshooting and problem-solving.	Medium
2.	Data Security Risks: Vulnerabilities in Firebase Firestore or OAuth authentication could compromise user data security.	Implement robust security measures such as encryption, access control, and regular security audits. Stay updated with security patches and best practices recommended by Firebase and OAuth documentation.	High
3.	Dependency Risks: Reliance on third-party services like Google Collab or Firebase Firestore may introduce dependency risks, such as service outages or changes in API.	Keep abreast of updates and announcements from service providers. Develop contingency plans to mitigate disruptions, such as alternative hosting solutions or local development environments.	Low
4.	Scalability Risks: Increased user demand may strain the scalability of the Flask server hosted on Google Collab, leading to performance	Monitor server performance metrics closely and implement auto-scaling or load balancing mechanisms as needed. Regularly assess server capacity and consider migrating to a more scalable hosting solution if required.	Low

	degradation or downtime.		
5.	Compliance Risks: Non-compliance with legal or regulatory requirements related to data privacy (e.g., GDPR) could result in legal consequences or reputational damage.	Ensure compliance with relevant data protection laws and regulations by implementing appropriate privacy policies, user consent mechanisms, and data handling practices. Conduct regular audits to ensure adherence to compliance standards.	High
6.	Mental Health Sensitivity Risks: Inadequate consideration of mental health sensitivity issues in the design and presentation of therapeutic images could potentially trigger or exacerbate emotional distress in vulnerable users.	Collaborate with mental health professionals or psychologists to ensure that the generated images adhere to best practices for mental health sensitivity and are inclusive and supportive of diverse emotional states. Incorporate features such as content warnings, user-controlled image customization options, and resources for seeking professional help or support. Establish protocols for monitoring user feedback and complaints related to emotional distress and promptly address any issues through appropriate interventions or referrals to mental health professionals.	High

Table 4.3.1: Risk Management

4.4 Functional Specifications

Following are the specific features and functionalities that the project includes.

1. Image Generation System:
    - The system generates therapeutic images using stable diffusion algorithms and deep learning techniques.
    - Users can specify preferences such as color palette, visual style, and theme to customize the generated images.
    - The system provides options for generating single images or batches of images based on user input.
2. User Feedback Mechanism:
    - Incorporated mechanisms for collecting user feedback on generated images, including ratings, comments, and preferences.
    - Utilized user feedback to continuously improve the image generation process and enhance the relevance and effectiveness of generated images.
3. Integration with Authentication and Database Systems:
    - Implemented user authentication using Firebase OAuth for secure access to the application and personalized user experiences.
    - Integrated Firebase Firestore as the database for storing user data, preferences, and image-related information.

4.4.1 Interfaces:

1. External connections needed
    - API endpoints are required to establish connections with external systems like Firebase. This enables data exchange and functionality integration between our application and these services, such as storing data in or retrieving images from Firebase.

    - Integration with app stores involves linking our application with platforms like Google Play Store for updates and deployments. This ensures that users can access the latest version of our app, receive updates seamlessly, and deploy new features or bug fixes efficiently

2. Internal connections needed

- Communication between the frontend and backend systems for data sharing.
- Interaction among modules in the project to ensure smooth operations.

3. Communication methods
    - RESTful API endpoints for seamless communication between components.
    - Data exchange protocols like JSON format for information transfer.

4. Graphical User Interface:

   Develop a user-friendly GUI for the mobile app, allowing users to interact with the image generation and recommendation systems. The interface must feature intuitive navigation and clear instructions for a seamless user experience.

4.4.2 Interactions

1. Performance and Scalability:
    - Ensured that the system is capable of handling concurrent user interactions and generating images efficiently, with minimal latency or downtime.
    - Implemented scalability measures such as load balancing and resource optimization to accommodate increasing user demand and workload.

2. Testing and Quality Assurance:
    - Conducted comprehensive testing of the image generation and recommendation systems to validate functionality, accuracy, and reliability.
    - Implemented automated testing procedures and quality assurance protocols to identify and address any bugs, errors, or performance issues proactively.

4.5 Testing and Feedback :

A comprehensive testing strategy was employed to validate the functionality, performance, and usability of the image generation and recommendation systems. This included conducting unit tests, integration tests, and system tests to verify the correctness of individual components and their interactions. Additionally, user acceptance testing (UAT) was conducted to assess the application's compliance with user requirements and expectations, soliciting feedback from a

diverse group of users representing different demographics, preferences, and mental health needs. Feedback mechanisms were integrated into the application to collect user input, ratings, comments, and preferences on generated images and overall user experience. This feedback was systematically analyzed and incorporated into iterative development cycles, driving continuous improvement and refinement of the image generation algorithms, recommendation system, and user interface design. Regular user testing sessions, surveys, and usability studies were conducted to gather qualitative and quantitative insights into user behavior, preferences, and satisfaction levels. The project team prioritized responsiveness to user feedback, promptly addressing issues, implementing requested features, and iteratively enhancing the application based on user input.

4.6 Security :

Security is paramount in the development and deployment of the "AI-Based Therapeutic Image Generation Using Stable Diffusion" project to safeguard user data, privacy, and system integrity. A multi-layered security approach will be adopted to mitigate potential threats and vulnerabilities across various aspects of the application. This includes implementing robust authentication mechanisms, such as Firebase OAuth, to ensure secure access to the application and protect user accounts from unauthorized access. User data, including preferences and feedback, will be encrypted both in transit and at rest using industry-standard encryption algorithms to prevent unauthorized interception or tampering. The application will adhere to best practices for secure coding and development, such as input validation, parameterized queries, and session management, to mitigate risks such as SQL injection, cross-site scripting (XSS), and session hijacking. Regular security audits and vulnerability assessments will be conducted to identify and remediate potential security weaknesses proactively. The infrastructure hosting the application, including servers, databases, and cloud services, will be configured with appropriate access controls, firewalls, and intrusion detection systems to defend against external attacks and unauthorized access attempts. Additionally, user education and awareness initiatives will be implemented to promote good security hygiene among users, including strong password practices, account monitoring, and reporting of suspicious activities. By prioritizing security throughout the development lifecycle and adopting a proactive stance towards risk management, the project aims to instill confidence in users and stakeholders regarding the confidentiality, integrity, and availability of their data and the overall security posture of the application.

Chapter 5
System Analysis Proposed Architecture

5.1 Design Consideration :

1. User Interface Design:

   The app's design aims to prioritize user-friendliness and intuitive navigation. It should present comprehensive information about electric vehicle details, including available charging spots, station locations, and pricing. Visual elements and layout are tailored to enhance user experience and facilitate quick access to relevant data, promoting seamless interaction with the application.

2. User Verification and Safety:

   Robust security measures are in place to ensure safe user interactions. This includes implementing secure login processes with encryption protocols for data protection. Access controls are used to manage user permissions and roles effectively, safeguarding sensitive information and maintaining the integrity of verification systems against unauthorized access.

3. Scalability:

   The system architecture is designed to scale efficiently in response to increasing demand. This involves optimizing performance metrics such as loading speeds and response times to support a growing user base and expanding the network of charging stations. Scalability ensures the platform remains responsive and capable of handling increased traffic without compromising user experience or system stability.

5.2 Assumptions and Dependencies:

5.2.1 Technical Assumptions and Dependencies

1. Availability of Resources: The project's dependencies include resources such as network connectivity to Google Cloud Platform (GCP) servers for access to the database as well as the backend services hosted on Flask through Google Colab, and mobile platforms (Android or iOS). These resources are critical for the functionality of

the app, and to ensure a seamless user experience.

2. Integration Dependencies: The system relies on integration with external services, primarily Firebase APIs, for critical functionalities like data storage using Firestore, image management via Firebase Storage, and user authentication through Firebase Auth. These integrations are vital for ensuring secure, efficient, and scalable backend operations, facilitating data management, content delivery, and user authentication.

5.2.2 Non-Technical Assumptions

3. Users are not younger than 18: We assume this to prevent exposure to potentially sensitive AI-generated content that could be triggering for younger audiences. Upon startup, users are informed of this restriction, ensuring transparency and responsible usage of the application's features and content, thereby complying with content regulations and ethical considerations, the application restricts access to users below 18 years old.
4. Users do not suffer from severe anxiety: While the app is designed to assist individuals with anxiety, it assumes that users do not have severe anxiety conditions. This is to ensure that the app's content and features are suitable and safe for users. During startup, a warning is issued to inform users of this assumption, emphasizing the app's intended audience and its limitations in supporting individuals with specific mental health needs. This approach promotes responsible usage and encourages users to seek appropriate professional support if they experience severe anxiety or related mental health issues beyond the app's scope.

5.3 General Constraints :

Time Constraints: The development process had to adhere to strict project deadlines, requiring efficient task management by the team. To overcome this, we effectively prioritize tasks, clearly communicate with each other, and implement agile methodologies to ensure timely delivery of project milestones. Proactive monitoring of progress and frequent status updates were essential to mitigate risks associated with time constraints and optimize resource allocation.

Resource Constraints: The limited availability of hardware and software resources significantly

impacted the speed and effectiveness of the development process. This constraint necessitated resource optimization, prioritization of critical tasks, and strategic outsourcing or collaboration to overcome limitations (see the use of GPUs and CUDA in section 8). Managing resource constraints required flexibility in resource allocation, and continuous evaluation of project requirements to ensure optimal utilization of available resources while maintaining productivity and quality standards.

Compatibility Challenges: Ensuring the app functioned well on various devices and platforms presented significant challenges during development. We conducted extensive testing across different setups, identified platform-specific issues, and employed adaptable design techniques to ensure consistent user experience across diverse environments. Collaboration with platform specialists and utilization of specialized testing tools aided in resolving compatibility issues, ensuring smooth operation on different devices and operating systems.

Testing Restrictions: We encountered limitations in testing resources and environments, which hindered comprehensive testing. To address this, we prioritized critical testing areas, leveraged automated testing tools, and simulated real-world scenarios as much as possible. Our development and testing teams collaborated closely to identify testing gaps and continuously refine our testing approaches. Despite these challenges, our goal was to deliver a robust and reliable product by regularly monitoring test outcomes and making ongoing enhancements to our testing methodologies.

5.4 Block Diagram

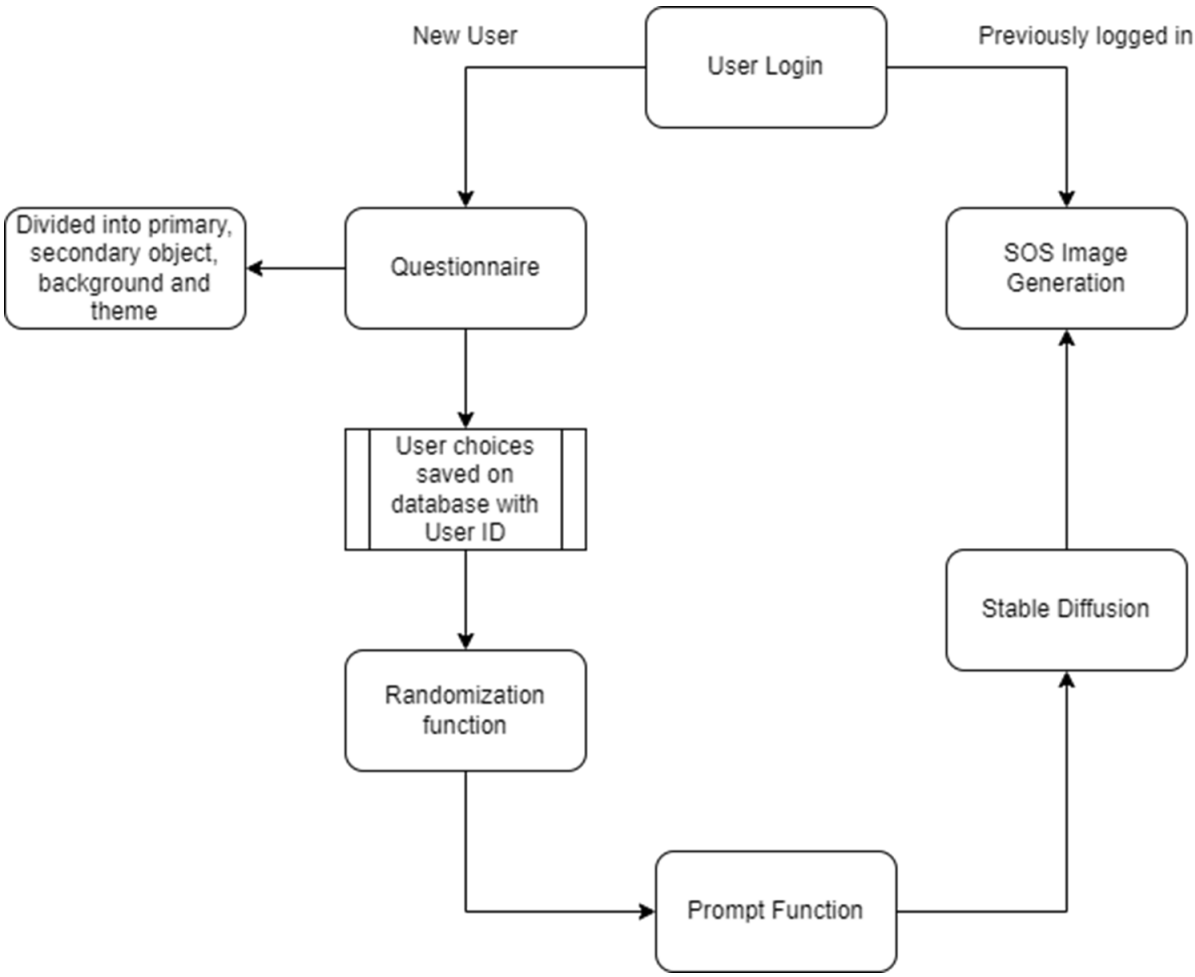

Figure 5.4.1 Block Diagram

5.5 System Architecture :

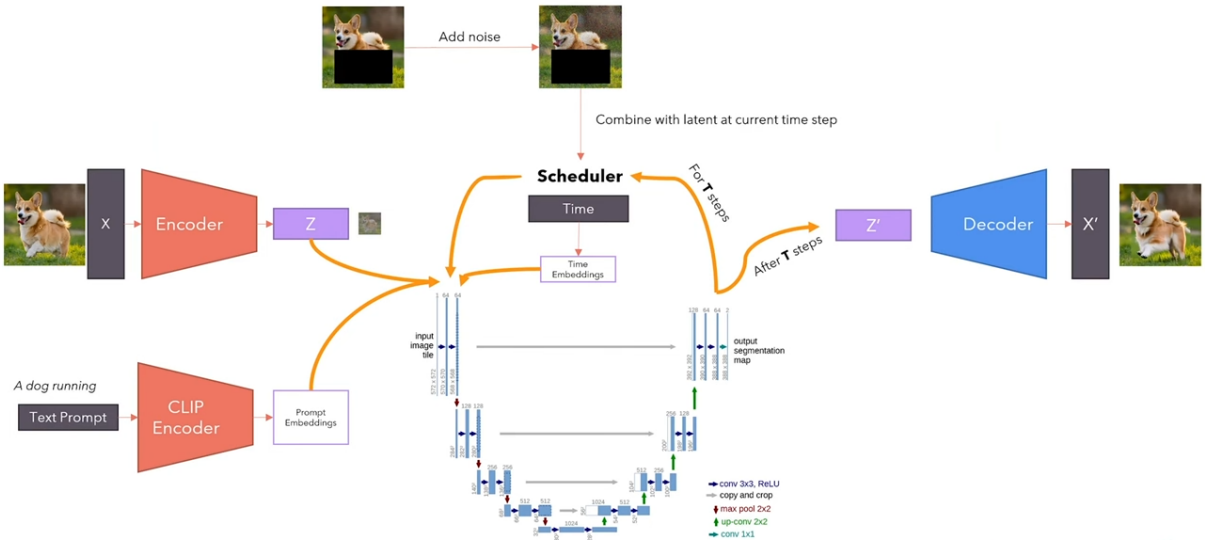

Figure 5.5.1 Architecture Diagram (Text to Image)

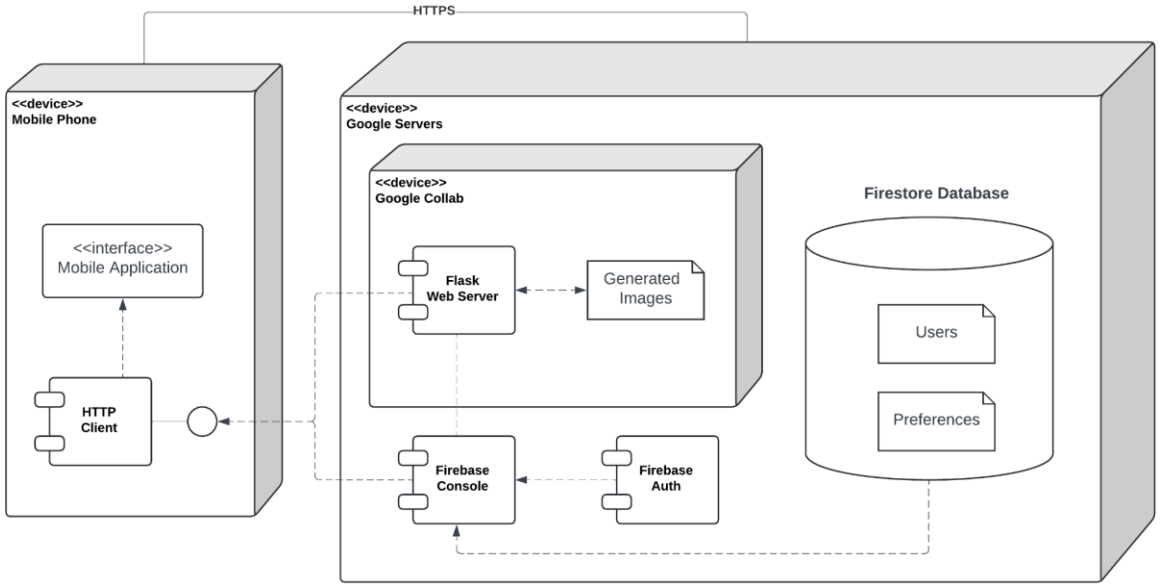

Figure 5.5.2 System Deployment Diagram

5.6 Modules of the Project :

1. User Recommendation System

- In-depth Research and Development: This module undertakes comprehensive research into reinforcement learning algorithms and user behavior analytics. The aim is to create a sophisticated recommendation engine that not only predicts but also adapts to user preferences in real-time.
- Continuous Updates and Improvement: Maintains an up-to-date understanding of the evolving landscape of recommendation systems. This involves assimilating the latest research and technological advancements to enhance the system's accuracy and efficiency.
- Integration and Feedback Loops: Offers critical feedback and insights, facilitating the seamless integration of the recommendation system with app development and image processing workflows. This ensures that the app's functionality aligns with user expectations and system capabilities.

2. Image Generation System

- Advanced Image Processing Implementation: Utilizes cutting-edge technologies such as Stable Diffusion or Generative Adversarial Networks (GANs) to process images. This module focuses on crafting high-quality images that are customized to match individual user preferences.
- Seamless Integration with Recommendation Systems: Incorporates sophisticated image processing capabilities within the recommendation framework, allowing for the delivery of personalized image suggestions to enhance user engagement.
- Utilization of Leading-Edge Techniques: Integrates the latest algorithms and state-of-the-art methods in image generation and processing, ensuring the production of visually appealing and contextually relevant images.

3. App Development

- Robust and Scalable Backend Architecture: Designs and develops the backend and API of the mobile application with a focus on scalability, security, and high performance. Ensures the architecture supports evolving user demands and complex data operations.
- Integration of Recommendation Model: Integrates the sophisticated recommendation model into the mobile app's framework, which enhances the user experience by ensuring smooth,

efficient data flow and responsive interaction.

- User-Centric Design Philosophy: Employs a design approach that is deeply rooted in user behavior research. This methodology informs the development of user-centric features and design elements, making the app intuitive and engaging for end-users.

4. Psychology Collaboration and Research/Data Processing

- Extensive Research and Analysis: Conducts a broad and thorough literature review on reinforcement learning, image processing techniques, and user recommendation systems. This research is critical in identifying and applying the most effective and innovative methodologies.
- Therapeutic Image Dataset Preparation: Engages in meticulous dataset preparation, curating and fine-tuning therapeutic images that are used to refine the image generation model. This dataset is instrumental in training the model to produce outcomes that are therapeutic and aligned with psychological research findings.

5.7 Low-level Design :

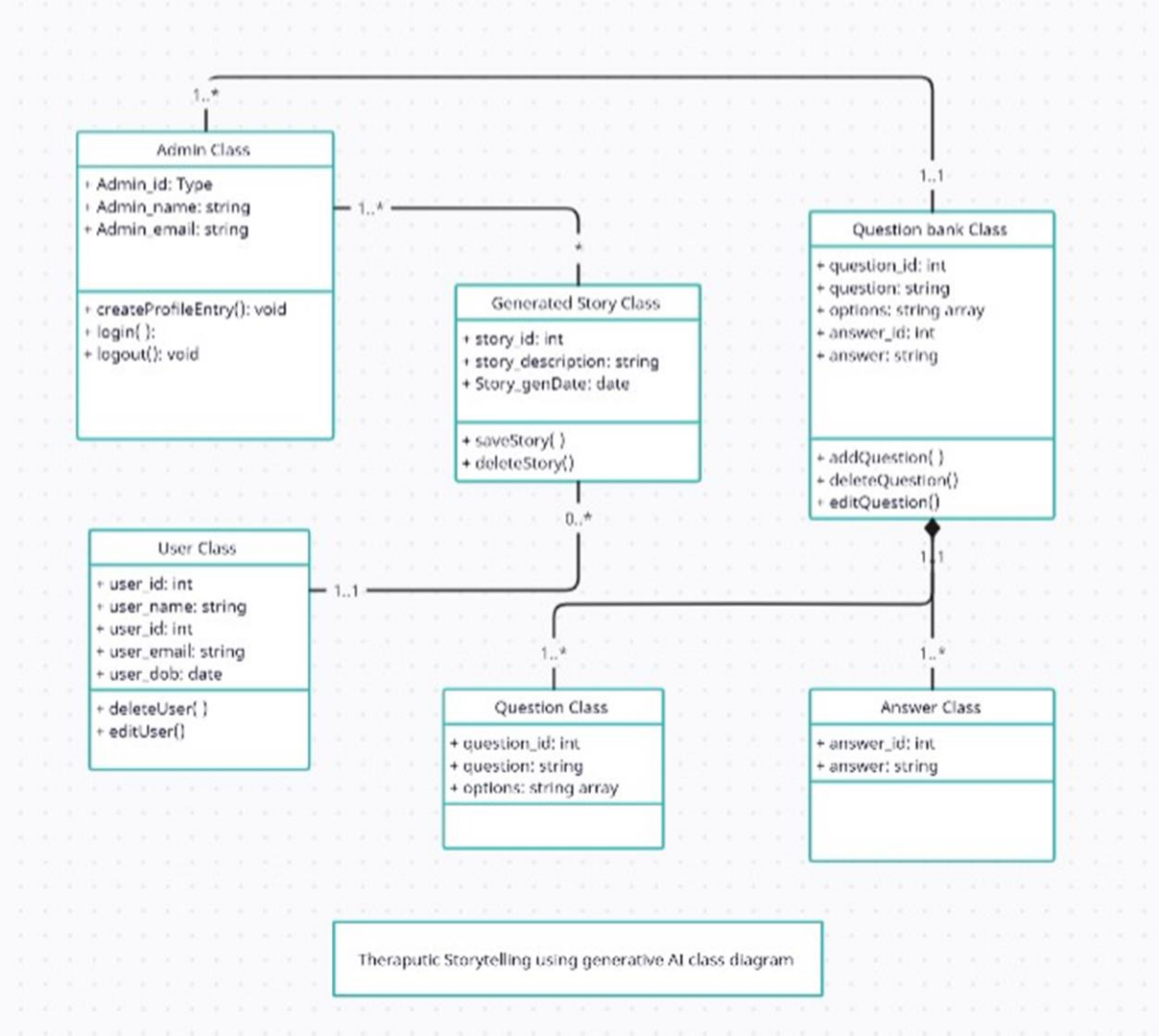

Figure 5.7.1: Class Diagram (Low-Level Design)

5.8 UML Diagrams :

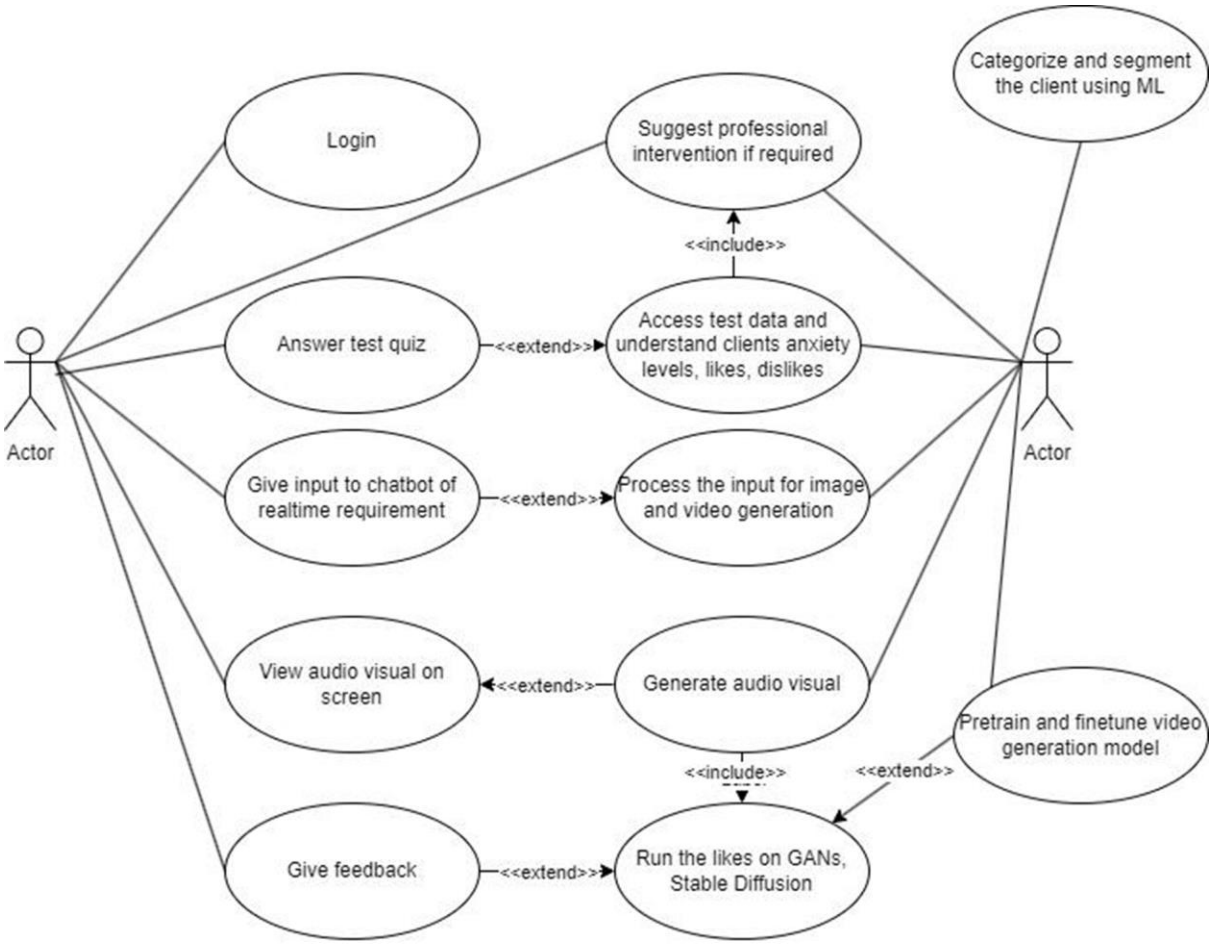

Figure 5.8.1: Use-Case Diagram

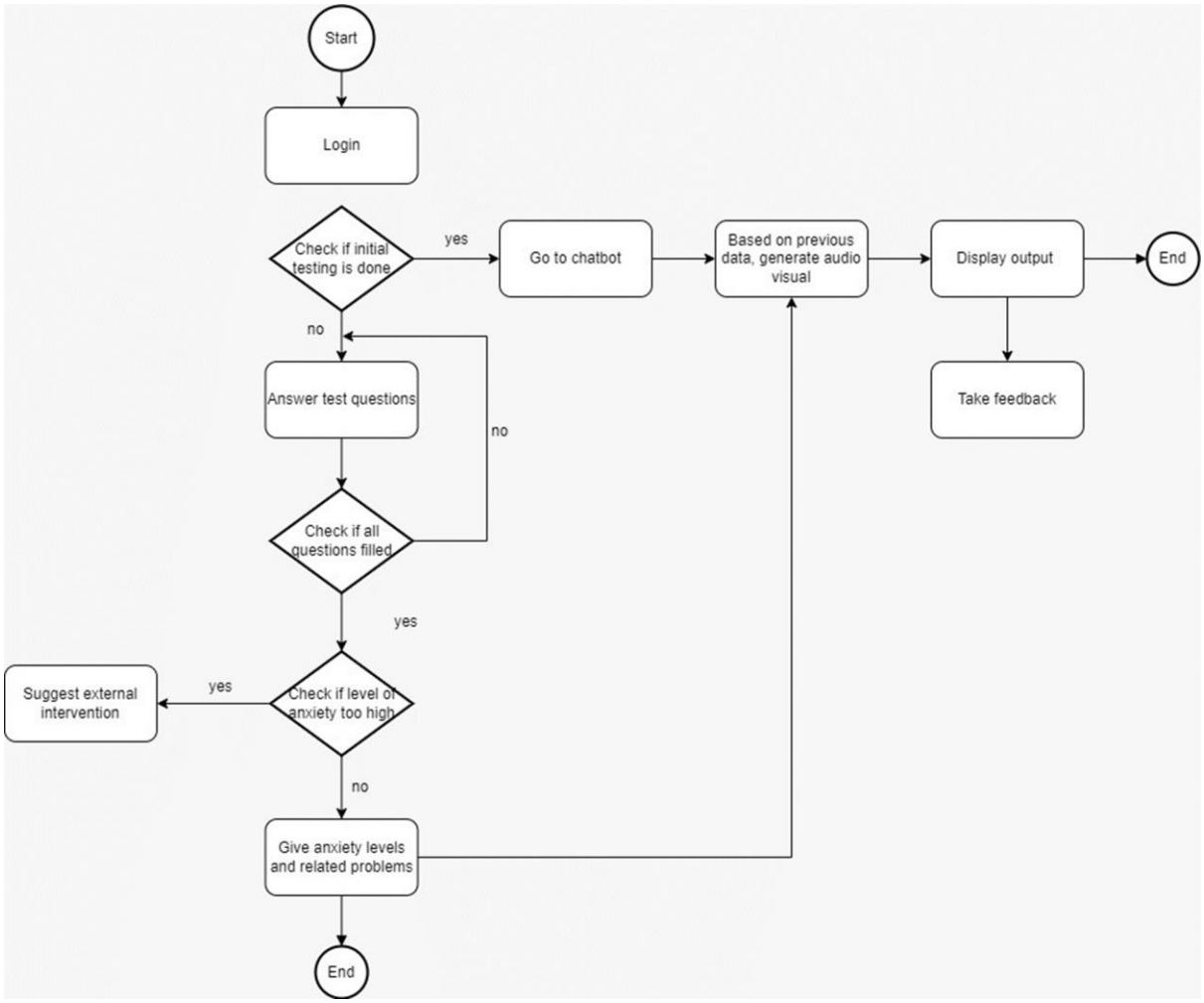

Figure 5.8.2 Activity Diagram

Chapter 6
Project Plan

6.1 Project Timeline :

Month	Task
March	• Conducted a Literature Survey to review existing research papers, articles, and studies related to stable diffusion algorithms, GANs, and therapeutic image generation. Summarized key findings and identified relevant methodologies and techniques for the project. • Gathered Requirements for running GANs/Stable Diffusion by consulting with domain experts, psychologists, and potential end-users to understand their needs, preferences, and expectations from the therapeutic image generation system. Documented the functional and nonfunctional requirements to guide the development process.
April	• Prepared datasets for training (images, user feedback data, etc.) by sourcing relevant image datasets from publicly available repositories and collecting user feedback data through surveys or interviews. Cleaned and preprocessed the datasets to ensure compatibility with the training algorithms and models. • Implemented Stable Diffusion for image processing using deep learning frameworks such as TensorFlow and PyTorch. Experimented with different model architectures, hyperparameters, and optimization techniques to achieve high-quality and emotionally resonant image generation results. • Integrated the recommendation model with the image processing module to create a unified system for generating therapeutic images based on user preferences and feedback. Developed APIs and communication protocols to facilitate seamless interaction between the recommendation and image processing components.

	• Started development of app backend and API to provide backend services for the mobile app, including user authentication, data storage, and interaction with the recommendation and image processing modules. Designed and implemented RESTful APIs for communication between the mobile app and backend server. • Finalized development of the recommendation model and image processing module by conducting thorough testing and validation. Validated the performance and accuracy of the models using benchmark datasets and real-world user feedback. Fine-tuned model parameters and algorithms based on validation results. • Tested integration between the recommendation model, image processing, and app backend to ensure smooth communication and interoperability between different system components. Conducted end-to-end testing scenarios to simulate user interactions and evaluate system behavior under various conditions. • Developed frontend of the mobile app to provide a user-friendly interface for accessing and interacting with the therapeutic image generation system. Designed intuitive user interfaces and implemented responsive layouts for seamless navigation and interaction on different devices and screen sizes. • Conducted testing and debugging to identify and address any issues or bugs in the application. Performed comprehensive testing across different devices, browsers, and operating systems to ensure compatibility and reliability.
May	• Deployed the app to the production environment by setting up hosting infrastructure, configuring servers, and deploying application code and resources. Monitored the deployment process and troubleshooted deployment-related issues and errors. • Conducted user acceptance testing (UAT) and gathered feedback from a diverse group of users representing different demographics, preferences, and mental health needs. Analyzed feedback to identify areas for improvement and refinement in the application. Incorporated user

	feedback into iterative development cycles to enhance user satisfaction and usability. • Wrote a research paper on the project to document the methodology, findings, and contributions of the project to the field of therapeutic image generation.

Chapter 7
Implementation

7.1 Methodology :

7.1.1. User Preferences and Identification:

1. User Identification and Preferences Tracking: Each user is uniquely identified via a User ID upon logging into the application. This ID facilitates personalized tracking and storage of user preferences and selections throughout their interaction with the application.
2. Image Generation Categories:
   - Primary Object: Users can choose a central object for their image from categories like animals, flowers, plants, and cartoons, which serve as the focal point.
   - Secondary Object: Complements the primary object by adding depth and context to the image.
   - Background: Options such as natural scenery, water bodies, abstract art, comforting scenes, and fantasy landscapes are offered to enhance the soothing effect of the image.
   - Theme: Users can select a color scheme that best suits their mood or preference, such as calming pastel tones or earthy shades, to set the overall ambiance of the image.

3. Example Questions for User Preferences:
   - Background Selection:
   "Which type of natural scenery do you prefer?"
   - Options: Forests, Mountains, Beaches, Lakes, None of the above, Other

   - Primary/Secondary Object Selection:
   "Which type of animal brings you more joy?"
   - Options: Dogs, Cats, Birds, Other (please specify)

   - Specific Preferences for Animal Type:
   If "Dogs" are selected:
   "Which type of dog appearance do you find most appealing?"
   - Options: Fluffy dogs, Sleek dogs, Small dogs, Large dogs, All of the above, Other (please specify)

   - Theme and Color Preferences:
   "If you had to choose a calming color palette, which would you prefer?"
   - Options: Soft pastel shades, Earthy tones, Other (please specify)

"Given your preference for soft and pastel colors, which shade do you find most soothing?"
- Options: Baby blue, Blush pink, Lavender, Mint green, Peach, Other (please specify)

7.1.2. Prompt Generation:

1. Dynamic Prompt Formulation: Utilizes a randomizer function to aggregate user-selected preferences across various categories to formulate a custom prompt. This structured prompt guides the image generation process by specifying the composition and theme of the desired image.

2. Structured Prompt Example:
   “Create a calming painting featuring a primary object of {primary object} and a secondary object of {secondary object}, set against a {background} with an overarching {theme} color scheme."

7.1.3. Image Generation:

1. Advanced Text-to-Image Conversion: Leverages the capabilities of Stable Diffusion, a state-of-the-art diffusion model, to transform text prompts into visually appealing images.
2. Iterative Image Refinement Process:
   - Forward Process: Introduces iterative noise to an initial image using a sequence of Gaussian distributions to explore diverse image possibilities.
   -Reverse Process: Employs a neural network to methodically reduce noise and refine the image towards the desired outcome as dictated by the user's prompt.
3. Integration of Prompt in Noise Correction:
   - A U-Net architecture is pivotal in adjusting noise levels, guided by a CLIP text encoder that aligns the noise adjustment process with the thematic elements of the prompt.
4. Efficient Image Handling with VAE:
   - Implements a Variational AutoEncoder (VAE) to efficiently compress and decompress image data, ensuring rapid processing without sacrificing image quality.
5. Comprehensive Workflow:
   - Commences with a text prompt that encapsulates user preferences.
   - Progresses by sampling random noise and compressing it, followed by iterative denoising aligned with the thematic prompt.
   - Culminates in a high-fidelity image that resonates with the user's aesthetic and emotional preferences.

7.2 Algorithm

```
function U-Net_Diffusion(input_image, iterations,
diffusion_coefficient)
    // Initialize U-Net architecture
    unet = initialize_unet()

    // Set initial image as input
    current_image = input_image

    // Loop through diffusion iterations
    for i = 1 to iterations do
        // Perform diffusion step
        diffused_image = apply_diffusion(current_image,
diffusion_coefficient)

        // Apply U-Net to enhance diffusion
        enhanced_diffused_image = unet(diffused_image)

        // Update current image
        current_image = enhanced_diffused_image
    end for

    return current_image
end function

function apply_diffusion(image, diffusion_coefficient)
    // Perform stable diffusion on the input image
    diffused_image = apply_diffusion_operator(image,
diffusion_coefficient)

    return diffused_image
end function
```

```
function apply_diffusion_operator(image, diffusion_coefficient)
    // Apply stable diffusion operator on the input image
    // Use appropriate numerical method (e.g., finite
differences)
    // to compute the diffusion equation
    // For example, you can use finite differences:
    // diffused_image = image + diffusion_coefficient *
laplacian(image)

    return diffused_image
end function

function initialize_unet()
    // Initialize U-Net architecture
    // This could involve defining the network architecture,
    // specifying layers, activation functions, etc.
    // You can use libraries like TensorFlow, PyTorch, etc.
    // to create and initialize the U-Net model.
    // For example, in PyTorch:
    // unet = U_Net()

    return unet
end function
```

7.3 Other Implementation Details:

1. The mobile application utilizes the Flutter framework developed by Google, which is a high-level SDK designed for building mobile applications using a declarative approach.

2. The app's technology stack integrates Firebase services for cloud-based storage of user preferences using Firestore, authentication through Firebase Auth, and image storage and delivery via Firestore Storage.

3. Google Colab, which runs on the Google Cloud Platform (GCP), acts as the foundational infrastructure for hosting the stable diffusion model. Additionally, GCP hosts the Flask web server that supports the app's backend functionality.

Chapter 8
Deployment Strategies

8.1 Different Deployment Strategies :

1. The Generative AI Model used for text-to-image generation is Stable Diffusion which is a state-of-the-art generative model based on diffusion models. The Python script consisted of 13 supporting modules, and a main driving module which imported the other modules and libraries. The total script length is aout 350-400 lines of code.
2. In order to efficiently develop and deploy such a big script, we used Visual Studio Code at first, which is a Development Environment provided by Microsoft. It was easy too implement the code, as we installed the required libraries. But there were some limitations as the code was running on the local device. Firstly, the processing power of CPU was limited and absence of GPU took the code above 30 mins to run and generate one image. This was a major drawback for our project.
3. Therefore, we resorted to using cloud-based computing platforms, where we could leverage the processing power of Graphical Processing Units to process the code and enhance the image generation process.
4. We deployed our model on the Google Colab Platform, which is a hosted Jupyter Notebook service that requires no setup to use and provides free access to computing resources, including GPUs and TPUs.
5. Colab provides free access to it's GPUs and has CUDA(originally short for Compute Unified Device Architecture) support which is a massively parallel architecture by NVIDIA. It also provides TPU which is used to process tensors.
6. We made use of T4 GPU for our purpose to achieve high-performance computing by making provisions in our code to use CUDA if available.
7. Running the code using T4 GPU significantly brought down our code runtime from 30 minutes to just 1 minute and we we're able to generate images in much faster way and the code became more deployable.
8. In google colab, every new runtime session, it's uploaded files are lost, due to security reasons, therefore we had to upload the supplementing modules every time. However,

the modules were not that big in size and could be easily uploaded. The problem was with the pre-trained weights and the merges.txt and vocab.json file, which was about 4 GB in size.

9. So we had to upload these files to the drive and then mount the drive in the google collab jupyter notebook, which we could access through the Content Files.
10. For hosting the code on a server, we made use of ngrok and flask. NGROK is a cross-platform application that enables developers to expose a local development server to the Internet with minimal effort. The software makes your locally-hosted web server appear to be hosted on a subdomain of ngrok.com.

Chapter 9

Results and Analysis

9.1 Application Interface :

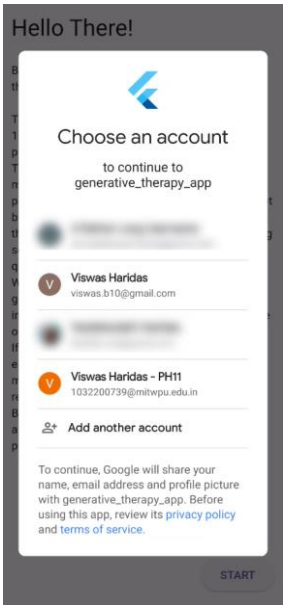

Figure 9.1.1 App UI - Login

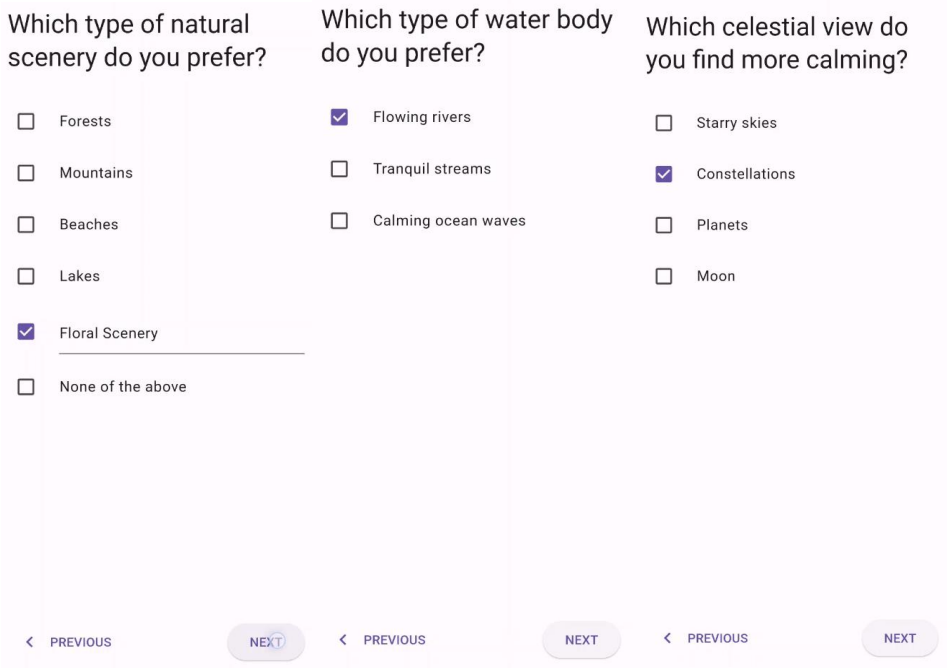

Figure 9.1.2: App UI - Questionnaire

Figure 9.1.3: Sample AI-Generated Therapeutic Images

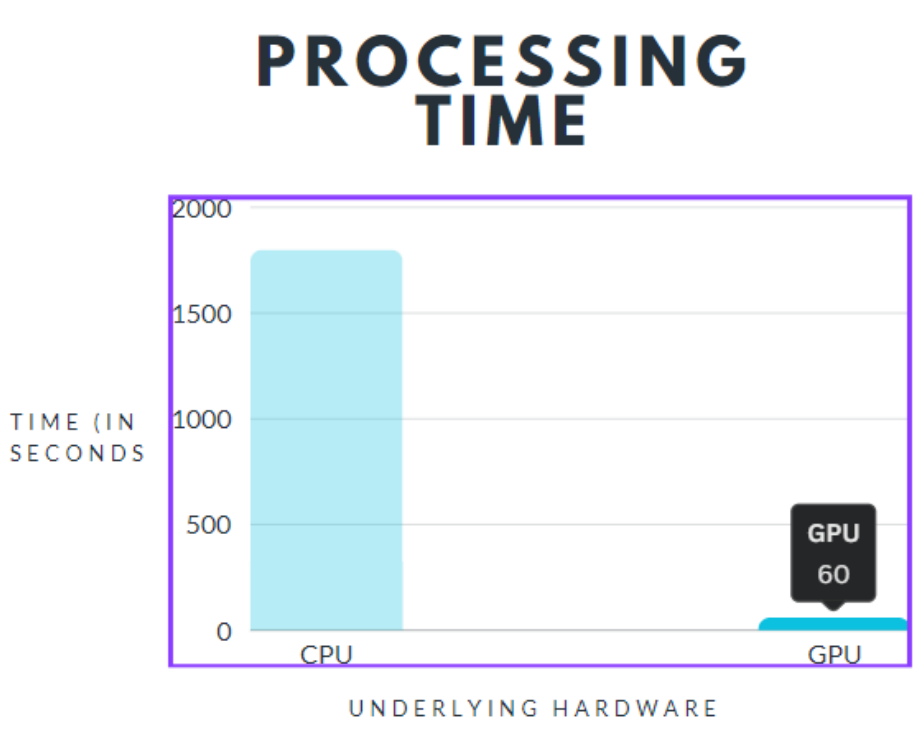

Figure 9.1.4: Execution Time difference between CPU and GPU

Applications

1. **Personalized Therapy Sessions:**
   The AI system can be used to deliver customized therapy sessions where users receive visual content tailored to their preferences and current emotional state. This could significantly enhance the effectiveness of therapy by ensuring that the interventions are closely aligned with the user's specific needs.

2. **Support for Mental Health Professionals:**
   This technology can serve as a tool for therapists and psychologists, providing them with additional resources to support their clients. The AI-generated content can be integrated into traditional therapy sessions, offering clients a novel way to engage with their treatment plans.

3. **Research and Development in Psychotherapy Techniques:**
   Beyond direct user interaction, the data collected and analysed by this AI system can provide valuable insights into the effectiveness of different therapeutic content and strategies. This could lead to new research opportunities and developments in psychotherapy practices, particularly in understanding how different stimuli can affect mood and anxiety levels.

4. **Enhanced Diagnostic Tools:**
   Can process vast amounts of patient data quickly, identifying patterns that may not be immediately apparent to human clinicians. This capability can help in the early detection of mental health disorders, allowing for timely intervention which is crucial for effective treatment.

5. **Crisis Intervention:**
   Monitor user behavior and language to detect signs of a mental health crisis, such as severe depression or suicidal thoughts. Once a potential crisis is identified, these systems can prompt immediate support interventions or alert medical professionals, thereby providing an additional safety net for individuals at risk.

6. **Automated Behavioral Recommendations:**

   AI can analyze patient data and daily activities to offer personalized lifestyle or behavioral changes that may benefit mental health. For example, suggesting specific mindfulness exercises or physical activities based on the individual's mood patterns and preferences.

7. **Training and Simulation for Professionals:**

   AI can create simulated environments or scenarios for training mental health professionals. These simulations can mimic real-life situations, allowing therapists, counselors, and psychologists to practice and refine their skills in a controlled but realistic setting.

8. **Longitudinal Mental Health Tracking:**

   AI systems can be used for ongoing monitoring of an individual's mental health status over long periods. This continuous data collection can help in understanding the long-term impacts of various treatment plans and in making necessary adjustments to enhance their effectiveness.

9. **Peer Support Optimization:**

   AI can facilitate the formation of peer support groups by matching individuals with similar conditions or therapy goals. This can enhance the support system for patients, providing them with shared experiences and tips for managing their conditions.

10. **Virtual Reality Therapy:**

    AI can be integrated with virtual reality (VR) to deliver immersive therapy experiences, such as exposure therapy for anxiety disorders or PTSD. AI-driven VR programs can adjust scenarios in real-time based on the patient's reactions and progress, providing a tailored therapeutic experience.

Conclusion

"AI-Based Therapeutic Image Generation Using Stable Diffusion" demonstrates a promising intersection of artificial intelligence and therapeutic practices aimed at enhancing mental health. By leveraging Stable diffusion models of AI, the project successfully developed a system capable of generating personalized therapeutic images tailored to individual psychological needs. This innovative approach not only provides immediate relief during acute anxiety episodes but also opens up new avenues for mental health support that are both adaptive and user-centered.

The data-driven insights derived from user interactions with the AI system could revolutionize the development of psychotherapy techniques, making mental health care more accessible and effective. The project's potential for future expansion and integration with other therapeutic modalities, such as guided meditation and breathing exercises, further underscores its versatility and capacity to address a broad spectrum of mental health challenges.

The project not only fulfills its aim of providing immediate therapeutic support through visually and emotionally resonant imagery but also contributes significantly to the ongoing discourse on the integration of technology in mental health care. It sets a benchmark for future projects in the realm of digital health therapies, highlighting the transformative potential of combining AI with traditional therapeutic practices.

Future Prospects of the Project

The existing scope of our project has several limitations. Firstly, the user -preferences for their liking and disliking is taken by answering a set of questions, but the number of questions is currently limited and covers only specific areas of choice. But they can be further extended by incorporating more detailed questions which will help in deriving user-specific results. Thereby, increasing the level of personalization of the application and making it deliver user-specific information. Secondly, the purpose of delivering results that will soothe or calm down a person is currently met only by generating the images as per the user preferences. But we can extend the therapy by also giving guided meditation or breathing sessions, which will further enhance the experience of a user. Thirdly, the scope of our application is limited to only instant and temporary therapy in times of emergency when access to a therapist is not available. Therefore, we also intend to put an assessment for understanding the stress levels of a person, then if the stress levels are too high, it is recommended that the person consult a therapist and get the proper treatment. For this reason, we have also provided a disclaimer stating the precautions and we will be providing the government-sponsored toll-free contact numbers for reaching out to the therapists who are available online also. We also plan to include mindfulness meditation in our app. Finally, we plan on including feedback from the user, which will further help us in improving the overall experience of our application. We also can perform data analysis which will further enhance the quality of output delivered by our application. The images generated are currently limited to general categories like Places, Animals, and other such objects, but we will try to generate images of different personas like Mother Teresa, Mahatma Gandhi, Abraham Lincoln, etc. which will further become more specific to user preferences.

References

[1] Clark, L. (2001). *SOS Help for Emotions: Managing Anxiety, Anger, and Depression* (illustrated, reprint). SOS Programs & Parents Press.

[2] Simon D'Alfonso, "AI in mental health", Current Opinion in Psychology, Volume 36, 2020, Pages 112-117, ISSN 2352-250X, https://doi.org/10.1016/j.copsyc.2020.04.005.

[3] S. Hamdoun, R. Monteleone, T. Bookman and K. Michael, "AI-Based and Digital Mental Health Apps: Balancing Need and Risk," in IEEE Technology and Society Magazine, vol. 42, no. 1, pp. 25-36, March 2023, doi: 10.1109/MTS.2023.3241309.

[4] Luxton, D. D. (Ed.). (2015). Artificial Intelligence in Behavioral and Mental Health Care (1st ed.). Academic Press.

[5] Kretschmar, T., & Tzschaschel, M. (2017). The Power of Inner Pictures: How Imagination Can Maintain Physical and Mental Health (1st ed.). Routledge. https://doi.org/10.4324/9780429482779

[6] Rusch, M. D., Grunert, B. K., Mendelsohn, R. A., & Smucker, M. R. (2000). Imagery rescripting for recurrent, distressing images. *Cognitive and Behavioral Practice*, 7(2), 173-182. https://doi.org/10.1016/S1077-7229(00)80028-0.

[7] Byrnes, S. R. (1996). *The Effect of Audio, Video, and Paired Audio-Video Stimuli on the Experience of Stress. Journal of Music Therapy, 33(4), 248–260.* doi:10.1093/jmt/33.4.248

[8] Bigham, E., McDannel, L., Luciano, I., & Salgado-Lopez, G. (2014). Effect of a Brief Guided Imagery on Stress. Biofeedback, 42(1), 28–35. doi:10.5298/1081-5937-42.1.07

[9] Wells, Deborah. (2005). The effect of videotapes of animals on cardiovascular responses to stress. Stress and Health. 21. 209 - 213. 10.1002/smi.1057.

[10] Ikei, H., Song, C., & Miyazaki, Y. (2023). Physiological adjustment effect of visual stimulation by fresh rose flowers on sympathetic nervous activity. *Frontiers in Psychology*, 14. https://doi.org/10.3389/fpsyg.2023.1159458.

[11]Divyanshu Mataghare, Shailendra S. Aote, Ramchand Hablani, "Text to Image Generation using Stable Diffusion," Eur. Chem. Bull., vol. 12, no. S3, pp. 4347-4353, 2023, doi: 10.31838/ecb/2023.12.s3.496.

[12]Zhou Zhou, Yunqing Zhu, Norihito Naka, "Text To Image Generation In DCGAN and Stable Diffusion Model," 2024, doi: 10.13140/RG.2.2.34276.96649.

[13]Rafael Hidalgo, Nesreen Salah, Rajiv Chandra Jetty, Anupama Jetty, Aparna S. Varde, "Personalizing Text-to-Image Diffusion Models," Journal, 2023.

Publication Details

Drafted a review paper, which has been submitted for a conference called “International Conference on Recent Trends in Computing”

AI-driven Mental Health Assessment and Personalized Therapeutic Storytelling

Vrishani Shah[1], Viswas Haridas[1], Yash Honrao[1], Prachiti Chitta[1], and Shamla Mantri[1]

[1] MIT World Peace University, Pune, India
{1032201102,1032200739,1032200191,1032200527,
shamla.mantri}@mitwpu.edu.in

Abstract. Sudden increase in mental health issues around the world calls for steps to be taken to help in providing easy and personalized access to therapy. With the advancements in the field of Generative AI and Neural Networks, the utilization of cutting-edge technology can help in making therapy more accessible and easier for individuals. We propose an application involving therapeutic audio-visuals which will act as an immediate step for individuals when they feel an anxiety attack kicking in. A combination of reinforcement learning to understand an individual's likes and dislikes along with their triggers will help the image generation model create personalized visuals which will help an individual during trying times. This review paper is a comparison of multiple methods that can be used to achieve this goal.

Keywords: Mental health, reinforcement learning, image generation, therapy.

1 Introduction

In recent times, the number of individuals suffering from some form of mental health illness symptoms has increased drastically [10]. Intervention is very important and it may not always be possible because it requires an individual to be consistent with their clinical gains. An online form of intervention specifically tailored to individuals is the need of the time. A combination of user interaction and cutting-edge technology is required to help people counter their mental distress.

Artificial intelligence (AI) can play an important role in increasing the awareness of the importance of mental health worldwide [11]. AI has proved useful in detecting ill-

Draft of the implementation paper has been completed.

AI-Based Therapeutic Image Generation Using Stable Diffusion

Vrishani Shah[1], Yash Honrao[1], Viswas Haridas[1], Prachiti Chitta[1], and Shamla Mantri[1]

[1] MIT World Peace University, Pune, India
vrishanishah20@gmail.com, yashhonrao2024@gmail.com,1032200739@mitwpu.edu.in,
prachitichitta@gmail.com, shamla.mantri@mitwpu.edu.in

Abstract

Due to the rise in mental health problems the need for digital solutions has increased, especially using Artificial Intelligence (AI). Audio-visuals have proven to have a positive impact on the human psyche. Certain audio-visuals have even proven to reduce stress and is a regular at health centers and hospitals. An emergency solution for people suffering from mental health issues that has an immediate effect is an important requirement. Using Stable Diffusion, an image generation model, this paper proposes generating personalized images of the users likes and creating a beautiful calming image which will help reduce anxiety and stress levels. Taking the users liking as particular categories and using a randomization function for selection of one choice from each category, an image is generated which consists of all the choices and creates an image which is easy on the eyes and can help the user calm down.

Keywords: Stable Diffusion, Image Generation, AI-Based Therapy

Plagiarism Report from any open source

turnitin

Similarity Report ID: oid:28480:56000226

PAPER NAME

springer-paper[1].docx

WORD COUNT

3196 Words

CHARACTER COUNT

18534 Characters

PAGE COUNT

8 Pages

FILE SIZE

97.8KB

SUBMISSION DATE

Apr 11, 2024 12:29 PM GMT+5:30

REPORT DATE

Apr 11, 2024 12:29 PM GMT+5:30

12% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 10% Internet database
- 8% Publications database
- Crossref database
- Crossref Posted Content database
- 3% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Small Matches (Less then 8 words)

Summary

AI-driven Mental Health Assessment and Personalized Therapeutic Storytelling

Vrishani Shah[1], Viswas Haridas[1], Yash Honrao[1], Prachiti Chitta[1], and Shamla Mantri[1]

[1] MIT World Peace University, Pune, India
{1032201102,1032200739,1032200191,1032200527,
shamla.mantri}@mitwpu.edu.in

Abstract. Sudden increase in mental health issues around the world calls for steps to be taken to help in providing easy and personalized access to therapy. With the advancements in the field of Generative AI and Neural Networks, the utilization of cutting-edge technology can help in making therapy more accessible and easier for individuals. We propose an application involving therapeutic audio-visuals which will act as an immediate step for individuals when they feel an anxiety attack kicking in. A combination of reinforcement learning to understand an individual's likes and dislikes along with their triggers will help the image generation model create personalized visuals which will help an individual during trying times. This review paper is a comparison of multiple methods that can be used to achieve this goal.

Keywords: Mental health, reinforcement learning, image generation, therapy.

1 Introduction

In recent times, the number of individuals suffering from some form of mental health illness symptoms has increased drastically [10]. Intervention is very important and it may not always be possible because it requires an individual to be consistent with their clinical gains. An online form of intervention specifically tailored to individuals is the need of the time. A combination of user interaction and cutting-edge technology is required to help people counter their mental distress.

Artificial intelligence (AI) can play an important role in increasing the awareness of the importance of mental health worldwide [11]. AI has proved useful in detecting illnesses with the help of professionals in the area. By using two techniques, digital phenotyping and natural language processing, AI can detect if a person is suffering from mental illnesses. AI's involvement may cause the dehumanization of the process and may cause problems. This is why having a personalized therapeutic method using AI is crucial. This technology holds both great promise to transform mental healthcare and its potential.

In this paper, we propose a combination of reinforcement learning and image generation to help people curb their anxiety levels. An individual can choose generated images having a therapeutic effect using which the model would learn the user's likes and

dislikes. An SOS button would help them generate therapeutic images when they feel an anxiety attack kicking in. This method is useful for low levels of anxiety when immediate help is required.

The model will generate certain images using certain prompts which fall into specific categories like animals, nature, shapes etc. The learning model will understand the likings of the user based on the choices of the user. A structure will be created to narrow down the likes and the image generation model will then take the output of the learning model as a prompt input and generate the appropriate image which will have a therapeutic effect on the user.

For image generation, we propose using Stable Diffusion and for RL, we propose a combination of natural language processing techniques and GANs. This is a review paper based on study of multiple, previously implemented methodologies which would apply to the proposed solution.

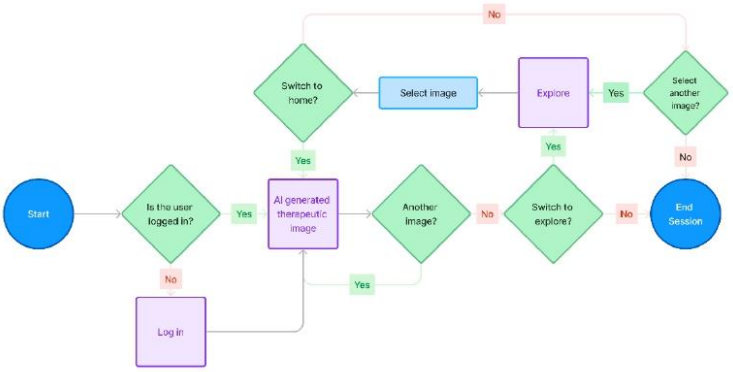

Fig. 1. User Flow Diagram of proposed method.

2 Literature Review

[1] The authors take the simple approach in a recommendation system, displaying k items and getting user feedback, the user would click or not click on the items, and a new set of items is then recommended. They talk about the importance of the user and how the user maximizes her rewards measures by choosing an item or not choosing an

item. Each step is taken according to the user's personal experience. They use a Generative Adversarial User Model which gives a probability distribution based on the reward function which gives an optimization function. The behavior model which mimics the actions of the user, acts as the generator and generates the users next step based on previous choices and the reward function acts as the discriminator and tries to differentiate the real user choices from the behavior model generated choices. They use a min-max function for optimization where the discriminator tries to increase the difference in statistics of the user and model actions and the generator tries to reduce that difference. They use a greedy method for giving the best choices to the user, namely W&D-LR, W&D-CCF and GAN-Greedy. They use online news platforms as their test set and try to find two outcomes, the reward and the CTR and number of iterations run. GAN-CDQN gives the best cumulative reward and cascading Q-networks works best for CTR. The combination of GAN and cascading Q-network for a combinatorial recommendation policy manages to handle abundant candidate items and give good results.

[2] The paper is performing a two stage assessment using deep reinforcement learning to find the aesthetic of images which are personalized to the user. Generally low-level visual features and high-level aesthetic attributes are proposed to identify the image aesthetics. They talk about deep CNNs which combine two images, using content of one image and style of another. By retouching the image with the user's preference, it is personalized according to the user. But due to the vast quantity of data required for this process, personalized image enhancement proves to be a difficult task. The author proposes User-Guided Personalized Image Aesthetic Assessment(UG-PIAA) which consists of two stages, namely user-guided image aesthetic ranking and personalized aesthetic distribution generation. They use two policy networks on a public aesthetic dataset and use the modified images by the user to modify and improve the policy networks. The images will be pushed to the image enhancement policy network and the image enhanced by the user will be the feedback for the network. This will then go to the ranking policy network and the users ranking will act as feedback for that policy network. A distribution is created with the output images of stage 1. A correlation coefficient is obtained and it corresponds to the aesthetic score and ranking of the images. This approach is better than the general black box retouching as it takes into account the users personal views. They use HE (Histogram Equalization) and DHE (Dynamic Histogram Equalization) to verify the results. The method proves to give excellent results and the generated aesthetic distribution is very close to the actual user aesthetic.

[3] In this paper, the authors train an RL-CycleGANs model for sim-to-real transfer. It is difficult to get the visual representations of objects with a task controller as it is very expensive, so they implement the policies on a simulation for training the Reinforcement Learning model. To address the simulation-to-reality gap, they use a General Adversarial Network(GAN). The RL-CycleGan helps in mapping a simulated image to the real image, then the final image is given to the model during testing. The CycleGAN losses help retain some of the input image and the RL-scene consistency helps in retaining the features which are important for the RL-trained Q-function. This helps train

policies with simulation data using domain adaptation techniques. The approach involves combining CycleGANs with a Q-learning model. The CycleGANs is used for mapping the simulation to the real images. The generator produces realistic images by reducing the adversarial loss function and the adversary tries maximizing it. Q-learning was used as a reinforcement learning technique. It learns a function that denotes the total reward, a loss is defined and the Q-function tries to minimize it. The CycleGANs perform much better than the general GANs. The overall project removes the need for task specific feature engineering and it mainly removes the visual gap in simulation and feature engineering.

[4] Diffusion Models is state-of-the-art technology for the synthesis of image data which make use of sequential denoising autoencoders for the picture generation process. But because these models work on pixel space, Diffusion Models sometimes need hours of GPU processing and inference is costly. This paper proposes a method by making use of latent space of potential pre-trained autoencoders which prepares Diffusion Models on obliged processing assets while keeping up with their performance. This paper suggests a method to train diffusion models by balancing between preservation of detail and complexity reduction. They have integrated cross-attention layers into the model architecture to convert diffusion models into powerful generators for text inputs which enables high-resolution synthesis in a convolutional manner. Since the real-world language and visual descriptions are chaotic and highly variable, the text-to-image techniques now in use create images while ignoring the distinction between foreground and background, which results in not clear images being generated. However, with the help of latent diffusion models (LDMs), the training and sampling effectiveness can be improved in Diffusion Models. LDMs require less computing resources than pixel-based methods but their sequential sampling process is slower than GANs. The proposed model shows a better inception score of 5.2 as compared to some other models like StackGAN , HDGAN etc.

[5] This paper identifies the problem with the Stable Diffusion Model which is popularly used for generating images from text prompts, that it is trained on a specific dataset which limits its ability to generate images outside of its data. They further propose a methodology to overcome this issue by incorporating two neural network based models – Hypernetworks and DreamBooth in order to allow generation of any image from the Stable Diffusion with less additional training. This paper also significantly emphasizes on subject personalization, which enables us to generate images for specific facets using a given reference set and the three models – Stable Diffusion, DreamBooth and HyperNetworks. The results show that there is a better degree of control with DreamBooth and it generates higher quality images in comparison to HyperNetworks. However, HyperNetwork does not require massive computing resources and can be potentially implemented on a typical computer. This project helps to generate versatile images with minimal training data and also adds real-life perspectives to the images.

[6] This paper implements both the DCGAN and Stable Diffusion Model on the MNIST dataset. It uses the CLIP ViT B/32 model for the text-embedding. It conducts a comparative study of both models by evaluating using the inception score model – inceptionv3. The comparative study puts forward the following conclusions – the learning speed of the Stable Diffusion is much faster than the DCGAN model. It is because the DCGAN has to spend time on generation as well as discrimination which is a binary classifier and this whole is a costly process. Moreover, the hyper-parameter tuning in the DCGAN is very challenging. Therefore, the DCGAN model is both quantitatively and qualitatively challenging. Whereas in the diffusion models, there are fewer hyper-parameters and therefore less time is spent on tuning these that enables the model to perform well. Their experiment shows that although Stable Diffusion is a more efficient way for generating images, it generates poor quality images. Whereas, the DCGAN can generate good quality images but it has difficulty in adjusting parameters. In order to improve the performance of the models, the paper suggests a method by combining the stable diffusion and DCGAN model advantages to implement the text-to-image model. Like integrating the discriminator of the DCGAN model to the stable diffusion model to improve the denoising process performance.

[7] Integrating machine learning (ML) and deep learning (DL) into mental health diagnostics enhances the accuracy and efficiency of diagnosing mental disorders, including schizophrenia, depression, anxiety, bipolar disorder, PTSD, anorexia nervosa, and ADHD. The literature utilizes systematic review methodologies like PRISMA to select studies that leverage ML and DL in mental health diagnostics and highlights the growing reliance on supervised learning, with Support Vector Machine (SVM) and Random Forest (RF) being predominant techniques due to their robust performance in classification tasks. In schizophrenia diagnosis, innovative approaches like Discriminant Autoencoder Networks (DANS) have shown high accuracy across different imaging sites, suggesting a consistent pattern recognition capability. Depression and anxiety diagnoses have also benefited from ML, with techniques such as sentiment analysis and hierarchical attention networks offering nuanced insights into patient conditions. For bipolar disorder, the fusion of structural and functional MRI data through SVM has improved diagnostic precision. The exploration of PTSD diagnosis through ML reflects a trend toward predictive modeling, utilizing large datasets to refine risk assessments. Anorexia nervosa research has seen the application of sentiment analysis in therapeutic contexts, enhancing the understanding of patient emotional states. ADHD diagnosis benefits from ML's ability to sift through extensive clinical records, improving identification accuracy.

[8] This paper delves into the application of Artificial Intelligence (AI) in psychological interventions and diagnosis, focusing on the transformative potential of machine learning (ML) and deep learning (DL) in the mental health domain. The integration of AI into psychological practices is explored through a comprehensive analysis of existing studies that highlight the effectiveness of AI tools in enhancing clinical outcomes. The paper elaborates on the utilization of AI in identifying and diagnosing various mental

disorders, emphasizing the role of DL in processing complex datasets to discern patterns and risk factors associated with mental health issues. It discusses how AI technologies, including sentiment analysis and predictive modeling, have been employed to advance the understanding and treatment of mental illnesses like depression, anxiety, and schizophrenia. The paper also examines the implications of AI in psychotherapeutic interventions, showcasing the potential of AI-assisted therapies to provide tailored and efficient treatment options. The discussion extends to the ethical and practical considerations of implementing AI in clinical settings, stressing the importance of safeguarding patient privacy and ensuring the accuracy of AI-powered diagnostic tools.

[9] The paper underscores the impact of machine learning (ML) and deep learning (DL) technologies in the realm of clinical psychology and psychiatry. These advanced computational methods are revolutionizing diagnostic, prognostic, and therapeutic processes, offering more precise, individualized, and predictive insights into mental health conditions. Significant attention is given to the application of ML in diagnosing psychiatric disorders helping in handling complex data sets to identify patterns and correlations that traditional statistical methods might overlook. This capability is particularly beneficial in differentiating between overlapping or comorbid conditions, thereby refining diagnostic precision and enabling more targeted therapeutic interventions. The review also delves into the predictive utility of ML, illustrating its potential in forecasting the trajectory of mental health conditions, which can significantly influence treatment planning and patient management. Furthermore, the research paper explores the integration of ML in treatment selection and optimization, showcasing its ability to analyze vast arrays of clinical and biological data to tailor treatment modalities to individual patient profiles. The literature review acknowledges the challenges and limitations of the need for larger, more diverse data sets to train the algorithms and the ethical considerations related to patient data privacy and algorithmic bias.

3 Results and Discussion

First, the person needs to be categorized into one of the few predefined categories which will be the deciding factor for which images to be generated in order to soothe the person. For the same purpose, we found that a Recommendation System will be suitable which is a subclass of Information filtering Systems that seeks to predict the rating or the preference a user might give to an item. So as a person chooses an image from multiple images, his chosen item's weights are updated in every iteration, until he gets categorized into one final subclass. There are several implementations of such Recommendation Models but we found User-Guided Personalized Image Aesthetic Assessment(UG-PIAA) to be most suitable for the task. The later part generates images related to the subclass that the person has been categorized in. For Image Generation purpose, there are multiple options available like RL-CycleGANs and Diffusion Models like Stable Diffusion. The RL-CycleGan helps in mapping a simulated image to the real image. The approach involves combining CycleGANs with a Q-learning model. Whereas, Diffusion Models is a state-of-the-art technology for the synthesis of image

data that make use of sequential denoising autoencoders for the picture generation process. The Stable Diffusion Model is popularly used for generating images from text prompts, but it is trained on a specific dataset which limits its ability to generate images outside of its data. The combination of AI and therapy promises an effective instant approach to solving mental disorders.

4 Conclusion

The escalating global demand for accessible mental health interventions emphasizes the urgency for innovative solutions. Our proposed application, leveraging Generative AI and Neural Networks, offers personalized therapeutic support for individuals grappling with anxiety. Through a synthesis of reinforcement learning and image generation techniques, we aim to provide immediate relief during anxiety episodes.

Our review shows the transformative potential of AI-driven interventions in mental healthcare. By integrating approaches such as User-Guided Personalized Image Aesthetic Assessment and RL-CycleGANs, we prioritize user preferences while tailoring therapeutic content. Additionally, our exploration of diffusion models, particularly the Stable Diffusion Model, highlights the importance of overcoming dataset limitations through innovative neural network-based models like Hypernetworks and DreamBooth.

The convergence of AI and therapy signifies a paradigm shift in mental health intervention, promising scalable and personalized solutions. While our framework represents a significant advancement, further research is imperative to unlock its full potential. By harnessing AI technologies, we can pave the way for a more accessible and user-centric approach to mental health support globally.

References

1. Chen, X., Li, S., Li, H., Jiang, S., Qi, Y. & Song, L. (2019). Generative Adversarial User Model for Reinforcement Learning Based Recommendation System. Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning Research 97:1052-1061 Available from https://proceedings.mlr.press/v97/chen19f.html.
2. Pei Lv, Jianqi Fan, Xixi Nie, Weiming Dong, Xiaoheng Jiang, Bing Zhou, Mingliang Xu, Changsheng Xu., "User-Guided Personalized Image Aesthetic Assessment Based on Deep Reinforcement Learning," in IEEE Transactions on Multimedia, vol. 25, pp. 736-749, 2023, doi: 10.1109/TMM.2021.3130752.
3. Kanishka Rao, Chris Harris, Alex Irpan, Sergey Levine, Julian Ibarz, Mohi Khansari; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 11157-11166
4. S. S. A. R. H. Divyanshu Mataghare, "TEXT TO IMAGE GENERATION USING STABLE DIFFUSION," European Chemical Bulletin, vol. 12, no. 3, pp. 4347 - 4353, June 2023.

5. N. S. R. C. J. A. J. A. V. Rafael Hidalgo, "Personalizing Text-to-Image Diffusion Models by Fine-Tuning Classification for AI Applications," in Conference: Intelligent Systems Conference (IntelliSys) 2023, Springer PublishersAt: Amsterdam, The Netherlands, Amsterdam, 2023.
6. Y. Z. N. N. Zhou Zhou, "Text To Image Generation In DCGAN and Stable Diffusion Model," New York, 2024.
7. Citation: N.K. Iyortsuun, S.-H. Kim, M. Jhon, H.-J. Yang, and S. Pant, "A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis," in Healthcare, vol. 11, no. 285, 2023, doi: 10.3390/healthcare11030285.
8. S. Zhou, J. Zhao, and L. Zhang, "Application of Artificial Intelligence on Psychological Interventions and Diagnosis: An Overview," Front. Psychiatry, vol. 13, article 811665, Mar. 2022, doi: 10.3389/fpsyt.2022.811665.
9. D. B. Dwyer, P. Falkai, and N. Koutsouleris, "Machine Learning Approaches for Clinical Psychology and Psychiatry," Annual Review of Clinical Psychology, vol. 14, pp. 91–118, 2018.
10. D'Alfonso S, Santesteban-Echarri O, Rice S, Wadley G, Lederman R, Miles C, Gleeson J and Alvarez-Jimenez M (2017) Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health. Front. Psychol. 8:796. doi: 10.3389/fpsyg.2017.00796
11. Minerva, F., Giubilini, A. Is AI the Future of Mental Healthcare?. Topoi 42, 809–817 (2023). https://doi.org/10.1007/s11245-023-09932-3

turnitin **Similarity Report ID:** oid:28480:56000226

12% Overall Similarity

Top sources found in the following databases:

- 10% Internet database
- Crossref database
- 3% Submitted Works database
- 8% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	**eurchembull.com** Internet	2%
2	**researchgate.net** Internet	2%
3	**arxiv.org** Internet	1%
4	**"Intelligent Systems and Applications", Springer Science and Business ...** Crossref	<1%
5	**University College London on 2022-04-10** Submitted works	<1%
6	**Pei Lv, Jianqi Fan, Xixi Nie, Weiming Dong, Xiaoheng Jiang, Bing Zhou, ...** Crossref	<1%
7	**Kanishka Rao, Chris Harris, Alex Irpan, Sergey Levine, Julian Ibarz, Mo...** Crossref	<1%
8	**Pei Lv, Jianqi Fan, Xixi Nie, Weiming Dong, Xiaoheng Jiang, Bing Zhou, ...** Crossref	<1%

Sources overview

9	deepai.org Internet	<1%
10	frontiersin.org Internet	<1%
11	Colorado Mesa University on 2023-12-11 Submitted works	<1%
12	University of Maryland, Global Campus on 2023-11-14 Submitted works	<1%
13	journals.sagepub.com Internet	<1%
14	ncbi.nlm.nih.gov Internet	<1%
15	"ECAI 2020", IOS Press, 2020 Crossref	<1%
16	link.springer.com Internet	<1%

PART B
Individual Contribution

A. YASH HONRAO

1) Name of the Student: Yash Milind Honrao

2) Module Title: Development of Generative Artificial Intelligence Model-Stable Diffusion, hosting it on the server and integrating it with front-end and database.

3) Project's Module Objectives -

1. The primary objective of this module was to take a prompt as input from the user and process it efficiently to deliver high-quality, high-resolution images as output.
2. The module should be able to efficiently fetch the user preferences from the Firebase database, therefore connecting it to the database.
3. Hosting it on the server, so that REST API Calls can be given from the front-end application.

4) Project's Module Scope -

The scope of this module is that it considers the user-preferences input by the users and generates the images that are to be displayed to the user on the front-end mobile application.

5) Project's Module(s) -

1. Development and deployment of Stable Diffusion using Python and Pytorch.
2. Python Flask and NGROK for hosting it on the server.
3. Firebase connectivity using the firebase module.

7) Module Interfaces

1. Connections between AI Model and the back-end Firebase database.
2. Connection between the model and the Front-end Mobile application.

8) Module Dependencies

1. Depends on the GPU(CUDA) for processing the text-to-image generation for faster processing time.
2. Depends on the internet connectivity for using Google Colab as the code is deployed on the cloud platform.

9) Module Design

The module consists of a CLIP Text Encoder for converting the text prompt to Text

embeddings, a Variational Auto Encoder for compressing and decompressing the pixels of the images, U-Net which is a convolutional neural network for predicting the amount of noise in the image and removing it.

10) Module Implementation

The module is implemented using Python, Pytorch and other modules which are all developed in Python

11) Module Testing Strategies

The testing was done by running the code and checking the amount of time required for running the code and also the quality of the image generated.

12) Module Deployment

The module was deployed on Google Colab which is a cloud-based platform for running Jupyter notebooks.

B. PRACHITI C CHITTA

1) Name of the Student: Prachiti C Chitta

2) Module Title: Integration of Psychology in Therapeutic Image Generation.

3) Project's Module Objectives - Individual Perspective:

1. Conducted comprehensive market research to understand the current landscape and implementation strategies of mindfulness applications. This involved analyzing the algorithms they employ, specifically focusing on how these applications utilize AI to enhance user engagement and therapeutic effectiveness.
2. Collaborated closely with the school of psychology to align the application's development with genuine psychological theories and user requirement insights, ensuring that the application addresses real-world needs effectively.

4) Project's Module Scope - Individual Perspective:

1. Executed an extensive literature review on stable diffusion algorithms, sophisticated image processing techniques, and advanced user recommendation systems to gather the latest methodologies that could be integrated into our project.
2. Implemented Firebase Firestore as the database to handle real-time data syncing and Firestore Storage.

5) Project's Module(s) - Individual Contribution:

1. Integration of Firestore database systems to ensure robust data management and user information security. Enhanced the application's responsiveness and data retrieval capabilities, which are pivotal for the user-centric features of the project.
2. Worked in tandem with the school of psychology to incorporate psychological insights into the app's design, ensuring that the image generation and recommendation systems are grounded in therapeutic efficacy.

6) Module Interfaces:

1. Ensured seamless integration between the user interface and the database, allowing for smooth interactions and an intuitive user experience.
2. Developed secure API endpoints for data exchange between the frontend and Firestore, ensuring data integrity and security.

7) Module Dependencies:

1. Relied on external libraries for reinforcement learning and image processing, integrating them with Firebase tools to support the application's functionalities.

2. Utilized collaborative platforms for continuous integration and deployment, ensuring that all components work seamlessly in unison.

8) Module Design:

Integrated advanced data management solutions using Firebase Firestore and Firestore Storage, enhancing real-time user data synchronization and secure content handling.

9) Module Implementation:

1. Customized reinforcement learning algorithms for user behavior prediction, enhancing the personalization of content.
2. Configured Firestore rules and storage schemas to optimize data storage and retrieval processes, critical for maintaining the application's performance.

10) Module Testing Strategies:

1. Led the efforts in unit testing of database functionalities and integration testing with the backend APIs to ensure robustness and reliability.
2. Coordinated with psychology experts to conduct user acceptance tests, ensuring the application’s features align with user expectations and therapeutic standards.

11) Module Deployment:

1. Managed the deployment of the database and storage solutions on cloud platforms-firebase, emphasizing scalability and accessibility.
2. Implemented advanced security configurations during deployment to protect user data and comply with data protection regulations.

C. VISWAS HARIDAS

1) Name of the Student: Viswas Haridas

2) Module Title: Design, Development, and Testing of Mobile Application

3) Project's Module Objectives - Individual Perspective

1. Establish the Flutter codebase for a fully functional mobile application.
2. Utilize Firebase for tasks such as user authentication, data storage and features like optimized file storage and delivery.
3. Build user interface elements and implement business logic for the app.
4. Ensure communication between Flutter and Firebase functionalities.

4) Project's Module Scope - Individual Perspective

This module emphasizes frontend development for apps using Flutter along with organizing code using the BLoC design pattern. It includes integrating Firebase to handle operations such as authentication and data handling.

5) Project's Module(s) - Individual Contribution

1. Setting up the Flutter framework and project layout.
2. Implementing BLoC design pattern for managing application state.
3. Incorporating Firebase services for tasks like authentication, data storage, and notifications.
4. Creating UI components that interact seamlessly with BLoC.

6) Hardware and Software requirements

Hardware: A computer running the Flutter SDK.

Software: Flutter SDK, An Integrated Development Environment (IDE), like Android Studio or Visual Studio Code, A Google account for Firebase, and a Web Browser

7) Module Interfaces

1. Connections between the Flutter app and the Firebase APIs for handling user authentication, data fetching, and persistent storage.
2. Links between Flutter's user interface elements and the business logic components (as separated by BLoC) for managing the local application state.

8) Module Dependencies

1. Utilization of Flutter SDK and plugins to integrate with Firebase.
2. Leveraging the BLoC library to control the state of the application.
3. Integration of Firebase SDK within the Flutter environment.

9) Module Design

1. Adherence to recommended project structure guidelines in Flutter.
2. Incorporation of the BLoC Architecture to segregate UI from business logic.
3. Modeling of the data structure in Firebase for user authentication and data storage.

10) Module Implementation

1. Initialization of a Flutter project using `flutter create`.
2. Installation of dependencies for Firebase integration and BLoC functionality.
3. Creation of authentication sequences using Firebase Auth.
4. Development of BLoCs for management of application states.
5. Building consistent UI using reusable components and themes.
6. Linking them with BLoC through `StreamBuilder`.

11) Module Testing Strategies

1. Conducting unit tests on BLoC logic utilizing `flutter_test`.
2. Executing integration tests on Firebase services.
3. Manual testing performed on UI components and user interactions to ensure end-to-end functionality.

12) Module Deployment

1. Rolling out the application onto test devices using Flutter's build tools.
2. Deployment of Firebase services for usage scenarios.
3. Establishment of continuous integration and deployment pipelines, for automated testing procedures.

D. VRISHANI SHAH

1) Name of the Student: Vrishani Shah

2) Module Title: Prompt Generation for Stable Diffusion, Surveying and Drafting Questionnaire and Research Paper Writing

3) Project's Module Objectives - Individual Perspective

1. To construct questions for user choices and develop prompt generation for the Stable Diffusion model.
2. To conduct thorough research on existing literature to understand recommendation model methodologies.
3. To investigate the advantages of using Stable Diffusion over other techniques like GAN, CLIP, etc.

4) Project's Module Scope - Individual Perspective

4. Constructing questions and prompt generation for the Stable Diffusion model.
5. Researching existing literature on recommendation models.
6. Exploring the pros of using Stable Diffusion compared to other techniques.

5) Project's Module(s) - Individual Contribution

1. Constructing questions and prompt generation for the Stable Diffusion model.
2. Researching existing literature on recommendation models.
3. Exploring the pros of using Stable Diffusion compared to other techniques.

7) Module Interfaces

1. Interaction with other modules involved in the recommendation system, such as data processing, model training, and user interface modules.

8) Module Dependencies

1. Data sources for constructing questions and prompts.
2. Existing literature on recommendation models for research and comparison purposes.
3. Integration with the Stable Diffusion model and other relevant modules.

9) Module Design

1. Designing a system to construct user choice questions and generate prompts for the Stable Diffusion model.
2. Designing a research strategy for exploring existing literature on recommendation models.
3. Planning experiments and analysis for comparing Stable Diffusion with other techniques.

10) Module Implementation

1. Implementing algorithms for constructing questions and generating prompts.
2. Implementing data collection and analysis methods for researching existing literature.
3. Implementing experiments to evaluate the effectiveness of Stable Diffusion.

11) Module Testing Strategies

1. Unit testing for question construction and prompt generation algorithms.
2. Evaluation of the research findings against established benchmarks and metrics.
3. Validation of the performance of Stable Diffusion through experiments and comparison with other techniques.

12) Module Deployment

1. Integration of the constructed questions and prompt generation system with the overall recommendation system.
2. Incorporating research findings into the development process of the recommendation model.
3. Deployment of Stable Diffusion as part of the recommendation system, if deemed effective.

Project to Outcome Mapping

Sr. No.	PRN No.	Student Name	Individual Project Student Specific Objective	Learning Outcomes mapped (To be filled by Guide)
1	1033220191	Yash Honrao	Developed, deployed, and tested the fully functional Stable Diffusion model for text-to-image generation. Successfully integrated the model with the Firebase database and fetched the user preferences. Hosted the model on the server and was able to accept Calls from the mobile applications and POST the output.	
2	1032200527	Prachiti C Chitta	Conducted market research on the approach and algorithms that mindfulness applications use. Collaborated with the school of psychology to understand user requirements. Conducted an extensive literature survey on reinforcement learning, image processing techniques, and user recommendation systems to identify relevant research and methodologies. Implemented Firestore database and Firestore storage.	

3	1032200739	Viswas Haridas	Designed, built, and tested the fully functional mobile app. Utilized Firebase for tasks such as user authentication, data storage and features such as optimized file storage and delivery. Integrated the stable diffusion model with the Flutter application	
4	1032201102	Vrishani Shah	Conduct thorough research on existing literature to develop an effective recommendation model. Constructed the questions for user choices and worked on the prompt generation for the Stable Diffusion model. Researched GAN, CLIP etc. which shows the pros of using Stable Diffusion.	

7% Overall Similarity

Top sources found in the following databases:

- 3% Internet database
- Crossref database
- 6% Submitted Works database
- 1% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	ahnashwin1305.medium.com Internet	<1%
2	assets.researchsquare.com Internet	<1%
3	University of Hertfordshire on 2023-08-28 Submitted works	<1%
4	Kaplan College on 2023-03-12 Submitted works	<1%
5	Wawasan Open University on 2024-03-15 Submitted works	<1%
6	Massey University on 2024-04-19 Submitted works	<1%
7	University of Technology, Sydney on 2023-10-29 Submitted works	<1%
8	Amity University on 2017-06-13 Submitted works	<1%

9	De Montfort University on 2024-05-01 Submitted works	<1%
10	De Montfort University on 2011-09-23 Submitted works	<1%
11	library.nexteinstein.org Internet	<1%
12	journal.iistr.org Internet	<1%
13	Virginia Community College System on 2021-07-01 Submitted works	<1%
14	Liberty University on 2024-05-07 Submitted works	<1%
15	University of Wales Swansea on 2024-04-15 Submitted works	<1%
16	University of Wales Institute, Cardiff on 2024-04-29 Submitted works	<1%
17	University of Wales Institute, Cardiff on 2024-05-02 Submitted works	<1%
18	Colorado State University, Global Campus on 2020-07-03 Submitted works	<1%
19	Asia Pacific University College of Technology and Innovation (UCTI) on... Submitted works	<1%
20	University of Technology, Sydney on 2020-10-24 Submitted works	<1%

21	bartleby.com Internet	<1%
22	ijaecs.iraj.in Internet	<1%
23	University of East London on 2023-05-07 Submitted works	<1%
24	ts2.space Internet	<1%
25	ouci.dntb.gov.ua Internet	<1%
26	Melbourne Institute of Technology on 2021-06-12 Submitted works	<1%
27	Osama Ahmed, Saleh Abdul Saleem, Aquib Ali Khan, Sana Daruwala, A... Crossref	<1%
28	RMIT University on 2023-06-08 Submitted works	<1%
29	University of Hertfordshire on 2023-08-28 Submitted works	<1%
30	University of Surrey on 2023-08-16 Submitted works	<1%
31	amazon.com Internet	<1%
32	aseanjournalofpsychiatry.org Internet	<1%

33	techdaddy.org Internet	<1%
34	University of Wollongong on 2023-12-07 Submitted works	<1%
35	tudr.thapar.edu:8080 Internet	<1%
36	tandfonline.com Internet	<1%
37	CTI Education Group on 2024-03-27 Submitted works	<1%
38	Study Group Australia on 2020-08-11 Submitted works	<1%
39	University of Sydney on 2022-07-12 Submitted works	<1%
40	University of Westminster on 2015-04-29 Submitted works	<1%
41	docplayer.net Internet	<1%
42	worldwidescience.org Internet	<1%
43	Brunel University on 2012-12-03 Submitted works	<1%
44	Deepali Javale, Govinda Talreja, Amey Somatkar, Neha Mohite, Mukta ... Crossref	<1%

45	Digvijay Patil, Shruti Hanchate, Shraavya R. Srinivasarao, Rajat Puri. "C... Crossref posted content	<1%
46	London Metropolitan University on 2023-07-21 Submitted works	<1%
47	Maastricht University on 2024-03-21 Submitted works	<1%
48	University of Hull on 2023-04-27 Submitted works	<1%
49	University of Sunderland on 2023-04-11 Submitted works	<1%
50	atharvacoe on 2024-01-31 Submitted works	<1%
51	pure.uva.nl Internet	<1%
52	jmir.org Internet	<1%
53	slideshare.net Internet	<1%