



Shri Vile Parle Kelavani Mandal's

DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)



SLIDESTER

Submitted in partial fulfillment of the requirements
of the course Innovative Product Development
under

T. Y. B. Tech. Artificial Intelligence and Machine Learning

By

HATIM MULLAJIWALA 60017220038

ANTARA KADAM 60018220099

HAADI RAKHANGI 60017220037

Guide

Mr. Nilesh Rathod

(PROFESSOR)

Mrs. Komal Patil

(PROFESSOR)

University of Mumbai

2022-2023

CERTIFICATE

This is to certify that the project entitled “**SlideSter**” is a bonafide work of **Haadi Rakhang**
(60017220037) submitted to the **Department of Artificial Intelligence and Machine Learning**
in partial fulfillment of the requirement for the course of Innovative Product Development.

(Prof. Nilesh Rathod)
Internal Guide

(Dr. Aruna Gawade)
Head of the Department

(Dr. Hari Vasudevan)
Principal

CERTIFICATE

This is to certify that the project entitled “**SlideSter**” is a bonafide work of **Antara Kadam (60018220099)** submitted to the **Department of Artificial Intelligence and Machine Learning** in partial fulfillment of the requirement for the course of Innovative Product Development.

(Prof. Nilesh Rathod)
Internal Guide

(Dr. Aruna Gawade)
Head of the Department

(Dr. Hari Vasudevan)
Principal

CERTIFICATE

This is to certify that the project entitled “**SlideSter**” is a bonafide work of **Hatim Mullajiwala (60017220038)** submitted to the **Department of Artificial Intelligence and Machine Learning** in partial fulfillment of the requirement for the course of Innovative Product Development.

(Prof. Nilesh Rathod)
Internal Guide

(Dr. Aruna Gawade)
Head of the Department

(Dr. Hari Vasudevan)
Principal

Project Report Approval for Innovative Product Development.

This project report entitled *of SLIDESTER* by **Haadi Rakhangi (60017220037)**, **Hatim Mullajiwala (60017220038)**, **Antara Kadam (60018220099)** is approved for the course of Innovative Product Development.

Examiners

1. Mr. Nilesh Rathod

2. Mr. Sudhir Dhekane

3. Mrs. Komal Patil

Date: 23/11/2023

Place: Mumbai

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Hatim Mullajiwala (60017220038)

Antara Kadam (60018220099)

Haadi Rakhanghi (60017220037)

Date : 28/11/2023

Abstract

Slidester, at the forefront of presentation innovation, amalgamates advanced Natural Language Processing (NLP), speech recognition, and generative AI to redefine the creation process. Leveraging Dynamic Slide Suggestions, it provides a structured inception by recommending slides based on input topics, streamlining initiation. The platform dynamically generates content-rich slides, enhancing visual appeal and informational depth. Image Generation further automates the process by fetching relevant visuals based on slide content. A Conversational Interface enables users to effortlessly modify, add, or extend slides through speech, incorporating a natural and intuitive editing experience.

One of Slidester's distinctive features is its ability to break free from the limitations of traditional Generative AI models, which are often "stuck in time" with information available only up to their training data. Slidester's generative AI model can access the latest information from the internet, ensuring that presentations are enriched with the most up-to-date and relevant data. Internet-Powered Content grants access to real-time data, ensuring presentations remain current. Slidester embraces Document Integration, allowing seamless inclusion of external data.

The provision of downloadable formats and creative templates adds versatility, making Slidester an indispensable tool for impactful, automated, and visually compelling presentations. The project's ultimate aim is to enhance productivity by automating the presentation creation process and eliminating time-consuming tasks. Students, professionals, and individuals seeking impactful presentations will find Slidester to be their ultimate companion. By employing NLP techniques, speech recognition, and image generation, Slidester sets out to redefine presentation creation as an intuitive, enjoyable, and efficient experience. Slidester offers a powerful tool that simplifies and expedites the process of creating presentations. Its seamless integration of NLP, speech recognition, image generation, and real-time internet data access ensures a user-friendly platform that empowers individuals to effortlessly curate well-structured and visually appealing presentations. Slidester marks a significant step forward in the world of presentation creation, promising to be the future of hassle-free and dynamic presentations.

Table of Contents

Chapter	Contents	Page No.
I	Introduction	I
	1.1 Aim of the Project	
	1.2 Motivation	
	1.3 Objective	
II	Literature Survey	II
III	Problem Description	III
	3.1 Problem Statement	
	3.2 Scope of the Project	
IV	Proposed Design	IV
VI	Results and Discussions	VI
VII	Conclusion	VII
VIII	References	IX

List of Figures

Fig. No.	Figure Caption	Page No.
1	Architecture Diagram	30

List of Abbreviations

Sr. No.	Abbreviation	Expanded form
1	PPT	Power Point Presentation
2	LCM - LoRA	Latent Consistency Model – Low Rank Adaptation
3	NLP	Natural Language Processing

1. Introduction

'SLIDESTER'-PowerPoint Presentation using NLP

In the dynamic landscape of presentations, Slidester stands as a trailblazer, orchestrating a paradigm shift in how narratives are crafted and communicated. As the nexus of technology and communication becomes increasingly vital, Slidester emerges as a transformative force, reshaping the very essence of the presentation creation process. This innovative platform seamlessly integrates advanced technologies, including Natural Language Processing (NLP), speech recognition, and generative AI, to redefine the user experience and elevate the art of presentation.

At the heart of Slidester's revolution is its pioneering feature: Dynamic Slide Suggestions. This groundbreaking innovation redefines the initiation phase, providing users with a structured starting point by recommending slides based on input topics. This not only streamlines the often daunting task of beginning a presentation but also reflects Slidester's commitment to understanding and anticipating user needs. It's more than a presentation tool; it's a creative companion that guides users through the inception of their presentation journey.

However, Slidester's prowess extends far beyond initiation. The platform dynamically generates content-rich slides, enhancing visual appeal while deepening the informational reservoir of each presentation. The infusion of Image Generation further automates the process, ensuring that relevant visuals seamlessly align with the content of each slide. This harmonious blend of technology and creativity marks Slidester as a transformative force where visual and textual elements coalesce effortlessly.

What propels Slidester into a league of its own is its liberation from temporal constraints. While traditional Generative AI models often remain "stuck in time," Slidester's generative AI model transcends these limitations by accessing the latest information from the internet. This ensures that presentations are not only visually engaging but also enriched with the most up-to-date and relevant data, reflecting the ever-evolving nature of the subject matter.

Slidester's commitment to real-time relevance is further exemplified through Internet-Powered Content, granting user access to live data. This ensures that presentations remain current, reflecting the

latest trends, statistics, and developments. The platform seamlessly embraces Document Integration, allowing for the inclusion of external data sources and reinforcing its adaptability to diverse information landscapes.

As Slidester strides towards redefining the fabric of presentation creation, it does so with the user's experience at its forefront. The provision of downloadable formats and creative templates adds a layer of versatility, making Slidester an indispensable tool for individuals across diverse domains. From students navigating the academic realm to professionals seeking impactful business presentations, Slidester is not merely a tool; it's a transformative force promising an intuitive, enjoyable, and efficient experience. It marks a significant leap forward, heralding the future of hassle-free and dynamic presentations in a technology-driven era.

1.1 Aim of the Project

The primary aim of the Slidester project is to revolutionize the landscape of presentation creation, transcending traditional methodologies and introducing a new era of efficiency, creativity, and user-centric design. At its core, Slidester seeks to redefine how narratives are crafted and communicated by seamlessly integrating advanced technologies, such as Natural Language Processing (NLP), speech recognition, and generative AI.

The cornerstone of this innovative endeavor is the development and implementation of Dynamic Slide Suggestions. This groundbreaking feature aims to streamline the initiation phase of presentation creation by providing users with a structured starting point based on input topics. By doing so, Slidester not only simplifies the often daunting task of starting a presentation but also demonstrates a commitment to understanding and anticipating user needs, positioning itself as a creative companion in the presentation journey.

Beyond initiation, Slidester endeavors to dynamically generate content-rich slides that enhance visual appeal and deepen the informational reservoir of each presentation. The infusion of Image Generation further automates the process, ensuring seamless alignment between relevant visuals and slide content. This harmonious blend of technology and creativity sets Slidester apart as a transformative force, where visual and textual elements coalesce effortlessly.

An essential aspect of the project's aim is to liberate Slidester from temporal constraints. Unlike traditional Generative AI models, Slidester's model accesses the latest information from the internet, ensuring that presentations are not only visually engaging but also enriched with the most up-to-date and relevant data, reflecting the ever-evolving nature of the subject matter.

Slidester also strives for real-time relevance through Internet-Powered Content, granting users access to live data. This commitment ensures that presentations remain current, reflecting the latest trends, statistics, and developments. The platform's embrace of Document Integration further reinforces its adaptability to diverse information landscapes, acknowledging the importance of incorporating external data sources.

As Slidester advances toward redefining the fabric of presentation creation, the user's experience

remains at the forefront. The provision of downloadable formats and creative templates adds a layer of versatility, making Slidester an indispensable tool for individuals across various domains. Whether it's students navigating the academic realm or professionals delivering impactful business presentations, Slidester aspires to be more than a tool—a transformative force promising an intuitive, enjoyable, and efficient experience, marking a significant leap forward in the technology-driven era of dynamic presentations.

1.2 **Motivation**

The motivation for the Slidester comes from our desire to create something revolutionary and unique. When our group was asked to pick a project topic, we wanted it to be fresh and not something that had been done before. We are really into AI and machine learning, so we were excited to work on a project that would make a difference.

We started looking for ideas online, but none of them really caught our attention. We wanted to work on something that could solve a real problem that people face in their everyday lives. Then, we had to make a presentation about our idea, and that's when it hit us - creating presentations is such a tedious and time-consuming task. We felt like we were wasting our time on it when we could be doing something more productive, like coming up with new ideas or studying things we love.

That's when we got the idea for "Slidester." We thought it would be amazing if we could create a system that takes away the tedious and time-consuming part of making presentations. With "Slidester," people can just speak their ideas, and the system will do the rest - generating slides, adding content, and even creating relevant images. It's all about making presentation creation easier and more enjoyable for everyone.

The motivation behind Slidester is to empower people to focus on the essence of their message and not get bogged down by the nitty-gritty of presentation-making. We want to give them a tool that saves time and makes the whole process more efficient. Whether you're a student, a professional, or anyone who needs to make presentations, Slidester aims to be the go-to solution for hassle-free and dynamic presentations.

1.3 Objective

The overarching objective of Slidester is to redefine the landscape of presentation creation, positioning itself as a transformative force that seamlessly integrates advanced technologies, such as Natural Language Processing (NLP), speech recognition, and generative AI. Building upon this foundation, our primary goal is to enhance the entire presentation creation process, making it intuitive, enjoyable, and efficient for users across diverse domains.

At the core of Slidester's mission is the aim to revolutionize the initiation phase through its groundbreaking feature: Dynamic Slide Suggestions. This innovation provides users with a structured starting point, recommending slides based on input topics and streamlining the often complex task of beginning a presentation. Slidester goes beyond being a mere tool; it aspires to be a creative companion that understands and anticipates user needs, guiding them through the inception of their presentation journey.

Expanding on this, Slidester seeks to dynamically generate content-rich slides, augmenting visual appeal and enriching the informational reservoir of each presentation. The integration of Image Generation further automates this process, ensuring seamless alignment between relevant visuals and the content of each slide. This harmonious fusion of technology and creativity sets Slidester apart as a transformative force where visual and textual elements coalesce effortlessly.

A distinctive objective is Slidester's liberation from temporal constraints. Unlike traditional Generative AI models, Slidester's model accesses the latest information from the internet, ensuring that presentations are not only visually engaging but also enriched with the most up-to-date and relevant data, reflecting the ever-evolving nature of the subject matter.

Slidester commits to real-time relevance through Internet-Powered Content, granting users access to live data. This ensures that presentations remain current, reflecting the latest trends, statistics, and developments. The platform's embrace of Document Integration further reinforces its adaptability to diverse information landscapes, acknowledging the importance of incorporating external data sources.

As Slidester strides towards redefining the fabric of presentation creation, it remains committed to enhancing the user experience. The provision of downloadable formats and creative templates adds a layer of versatility, making Slidester an indispensable tool for individuals across diverse domains. Whether navigating the academic realm or delivering impactful business presentations, Slidester is not just a tool but a transformative force promising an intuitive, enjoyable, and efficient experience, marking a significant leap forward in the technology-driven era of dynamic presentations.

II. Literature Survey

In a paper called 'Development Of Gui For Text-To-Speech Recognition Using Nlp' from [1] tells us that Natural language processing is a widely used technique by which systems can understand the instructions for manipulating text or speech. In the present paper, a Text-to- speech synthesizer is developed that converts text into spoken word, by analysing and processing it using NLP and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text. Here we developed a useful text-to- speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3 file.

A research paper from [2] tells us about text conversion. A real time speech to text conversion system converts the spoken words into text form exactly in the similar way that the user pronounces. We created a real time speech recognition system that was tested in real time noiseous environment. We used the design of a bidirectional nonstationary Kalman filter to enhance the ability of this Real time speech recognition system. Real time speech to text conversion system introduces conversion of the uttered words instantly after the utterance. speech recognition system. We have used our own created database for its flexibility and TIDIGIT database for its accuracy comparison with the HMM based speech recognition system. MFCC features of speech sample were calculated and words were distinguished according to the feature matching of each sampled word.

The research paper [3] delves into the implementation of an audio-visual summarizer that achieves the aforementioned motive. With audio evidence on the speaker's delivery, paired with visual images of PowerPoint slides or handwritten material that is presented in the seminars, this device provides a smart solution of summarizing the entire presentation and logging the summary to a remote database server from where it is accessed through a user-end software application. The prototype comprises a Raspberry Pi coupled with a camera and a microphone. The prototype uses a fast RCNN model for text detection, Open Source Computer Vision for text extraction, Google Speech Recognition and NLP concepts for generating the summarized data.

All the research given in [4] tells about a deep learning-based architecture for semantically consistent image generation: recurrent convolutional generative adversarial network (RC-GAN). RC-GAN successfully bridges the advancements in text and picture modelling, converting visual notions from words to pixels. The proposed model was trained on the Oxford-102 flowers dataset, and its performance was evaluated using an inception score and PSNR. The experimental results demonstrate

that our model is capable of generating more realistic photos of flowers from given captions, with an inception score of 4.15 and a PSNR value of 30.12 dB, respectively. In the future, we aim to train the proposed model on multiple datasets.

Paper from [5] tell that Natural language processing is a widely used technique by which systems can understand the instructions for manipulating text or speech. In the present paper, a Text-to- speech synthesizer is developed that converts text into spoken word, by analysing and processing it using NLP and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text. Here we developed a useful text-to- speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3 file.

Paper from [6] proposed a system with an idea of automatically generate presentation slides from textual data. The slides mostly incorporate of only the important points related to the topic. There are various tools and systems are available in the market which only address with formatting of the slides but not the content. This paper proposes an idea of automatically ppt generation from text. This will finally help in reducing a great amount of the presenter's time and efforts. The intended system works on NLP rules to classify data for the desired slides.

Paper from [7] show an end-to-end deep learning approach can be used to recognize either English or Mandarin Chinese speech—two vastly different languages. Because it replaces entire pipelines of hand-engineered components with neural networks, end-to-end learning allows us to handle a diverse variety of speech including noisy environments, accents and different languages. Key to our approach is our application of HPC techniques, resulting in a 7x speedup over our previous system. Hence, using a technique called Batch Dispatch with GPUs in the data center, we show that our system can be inexpensively deployed in an online setting, delivering low latency when serving users at scale.

The paper published in 2022 [8] Imagen builds on the power of large transformer language models in understanding text and hinges on the strength of diffusion models in high-fidelity image generation. pretrained on text-only corpora, are surprisingly effective at encoding text for image synthesis. To assess text-to-image models in greater depth, we introduce DrawBench, we compare Imagen with recent methods including and find that human raters prefer Imagen over other models in side-by-side comparisons, both in terms of sample quality and image-text alignment.

SR NO.	PAPER NAME -TITLE	PUBLICATIONS	AUTHOR/YEAR	FINDINGS
1	DevelopmentOf Gui For Text-To-Speech Recognition Using Nlp	IEEE	Partha Mukherjee, Soumen Santra, Subhajit Bhowmick, Ananya Paul, Pubali Chatterjee, Arpan Deyasi	<p>Natural language processing is a widely used technique by which systems can understand the instructions for manipulatingtext or speech. In the present paper, a Text-to- speech synthesizer is developed that converts text into spoken word, by analysingand processing it using NLP and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text.</p> <p>Here we developed a useful text-to- speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3 file.</p>
2	A Real Time Speech To Text Conversion System Using Bidirectional Kalman Filter In Matlab	ICACCI	Neha Sharma Shipra Sardana	<p>A real time speech to text conversion system converts the spoken words into textform exactly in the similar way that the userpronounces. We created a real time speechrecognition system that was tested in real time noiseous environment. We used the design of a bidirectional nonstationary Kalman filter to enhance the ability of this Real time speech recognition system. Real time speech to text conversion system introduces conversion of the uttered words instantly after the utterance. speech recognition system. We have used our owncreated database for its flexibility and TIDIGIT database for its accuracy comparison with the HMM based speech recognition system. MFCC features of speech sample were calculated and words were distinguished according to the feature matching of each sampled word.</p>

3	A Cost-Effective Audio-Visual Summarizer for Summarization of Presentations and Seminars	IEEE	Anmol Bhat, Aneesh C Rao, Anirudh Bhaskar, Adithya V, Prof. Pratiba D	<p>This paper delves into the implementation of an audio-visual summarizer that achieves the aforementioned motive. With audio evidence on the speaker's delivery, paired with visual images of PowerPoint slides or handwritten material that is presented in these seminars, this device provides a smart solution of summarizing the entire presentation and logging the summary to a remote database server from where it is accessed through a user-end software application. The prototype comprises a Raspberry Pi coupled with a camera and a microphone. The prototype uses a fast RCNN model for text detection, Open Source Computer Vision for text extraction, Google Speech Recognition and NLP concepts for generating the summarized data.</p>
4	Text-to-Image Generation Using Deep Learning	MDPI	Sadia Ramzan, Muhammad Munwar Iqbal, Tehmina Kalsum.	<p>A deep learning-based architecture for semantically consistent image generation: recurrent convolutional generative adversarial network (RC-GAN). RC-GAN successfully bridges the advancements in text and picture modelling, converting visual notions from words to pixels. The proposed model was trained on the Oxford-102 flowers dataset, and its performance was evaluated using an inception score and PSNR. The experimental results demonstrate that our model is capable of generating more realistic photos of flowers from given captions, with an inception score of 4.15 and a PSNR value of 30.12 dB, respectively. In the future, we aim to train the proposed model on multiple datasets.</p>

5	Content Based Internet Search Engine For Analysis And Archival Of Mpeg-1 Compressed Newsfeeds	IEEE	Odej K.40 , Gerhard R. Joubert	<p>Natural language processing is a widely used technique by which systems can understand the instructions for manipulating text or speech. In the present paper, a Text-to- speech synthesizer is developed that converts text into spoken word, by analysing and processing it using NLP and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text.</p> <p>Here we developed a useful text-to- speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3 file.</p>
6	Enhancing Automatic PPT Generation Technique through NLP for Textual Data	IJARCCE	Pooja Belote, Sonali Bidwai, Snehal Jadhav, Pradnya Kapadnis ⁴ , Nakul Sharma ⁵	<p>Proposed system come with an idea of automatically generate presentation slides from textual data. The slides mostly incorporate of only the important points related to the topic. There are various tools and systems are available in the market which only address with formatting of the slides but not the content. But it having performance issues regarding of sentence extraction, so this paper proposes an idea of automatically ppt generation from text. This will finally help in reducing a great amount of the presenter's time and efforts. The intended system works on NLP rules to classify data for the desired slides.</p>

7	End-to-End Speech Recognition in English and Mandarin	IEEE	Dario Amodei, Rishita Anubhai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, Jingdong Chen, Mike Chrzanowski, Adam Coates, Greg Diamos, ErichElsen, Jesse Engel, Linxi Fan, Christopher Fougner, Tony Han, Awni Hannun, Billy Jun, Patrick LeGresley, Libby Lin, Sharan Narang	<p>We show that an end-to-end deep learning approach can be used to recognize either English or Mandarin Chinese speech—two vastly different languages. Because it replaces entire pipelines of hand- engineered components with neural networks, end-to-end learning allows us to handle a diverse variety of speech including noisy environments, accents and different languages. Key to our approach is our application of HPC techniques, resulting in a7x speedup over our previous system.</p> <p>Hence, using a technique called Batch Dispatch with GPUs in the data center, we show that our system can be inexpensively deployed in an online setting, delivering lowlatency when serving users at scale.</p>
8	Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding	Google Research, Brain Team	Chitwan Saharia , William Chan, Saurabh Saxena , Lala Li†, Rapha Gontijo Lopes, Tim Salimans, Jonathan Ho , David J Fleet , Mohammad Norouzi.	<p>Imagen builds on the power of large transformer language models in understanding text and hinges on the strength of diffusion models in high-fidelity image generation. Our key discovery is thatgeneric large language models,pretrained on text-only corpora, are surprisingly effective at encoding text for image synthesis. To assess text-to-image models in greater depth, we introduce DrawBench, we compare Imagen with recent methods including and find that human raters prefer Imagen over other models in side-byside comparisons, both in terms of sample quality and image-text alignment.</p>

Survey Outcome:

The literature survey conducted for the Slidester project reveals several key findings and existing works in the field of natural language processing (NLP), speech recognition, and image generation, which are crucial components of the project.

The survey highlights the development of text-to-speech synthesizers using NLP and digital signal processing technology to convert text into synthesized speech. This approach provides a foundation for Slidester's voice input feature, where users can speak slide titles to generate dynamic presentations. Real-time speech-to-text conversion systems are also explored, which can be applied to Slidester's speech recognition capabilities, enabling users to input content effortlessly. The survey discusses the use of deep learning-based architectures, such as recurrent convolutional generative adversarial networks (RC-GAN), for image generation. These models can be employed in Slidester's image generation feature, which automatically creates relevant images for each slide's content, enhancing visual appeal.

Additionally, the survey mentions approaches to automatic presentation slide generation from textual data, which aligns with Slidester's aim to streamline the presentation creation process. Furthermore, the literature survey discusses end-to-end deep learning approaches for speech recognition, making it possible to handle diverse speech languages and environments. This is pertinent to Slidester's goal of integrating speech recognition technology.

Lastly, the survey highlights the effectiveness of large transformer language models in understanding text and their application in text-to-image synthesis. This aligns with Slidester's utilization of NLP techniques to generate presentation slides from text data efficiently.

III. Problem Statement:

The traditional methods of creating presentations pose significant challenges due to their time-consuming and tedious nature. Users often find themselves sifting through vast amounts of information, searching for suitable images, and selecting appropriate templates, which hinders productivity and creativity. This labor-intensive process can lead to frustration and a misallocation of valuable time that could be better utilized for more meaningful tasks. Additionally, manual data entry for each slide can be error-prone, resulting in suboptimal presentation quality.

The lack of efficient tools that streamline the presentation creation process leaves users grappling with the complexity of organizing content, selecting visuals, and maintaining audience engagement. The absence of an automated and user-friendly platform results in a persistent problem for individuals from various domains, including students, educators, professionals, and entrepreneurs.

3.2 Scope of the Project

The scope of the Slidester project is expansive and forward-thinking, with a core focus on revolutionizing the presentation creation process through the seamless integration of advanced technologies. This comprehensive and user-friendly platform aspires to cater to the diverse needs of individuals across various domains, setting itself as a transformative force in the landscape of dynamic presentations.

Central to the project's scope is the groundbreaking feature of Dynamic Slide Suggestions, redefining the initiation phase by recommending slides based on user input topics. This innovative approach streamlines the often complex task of beginning a presentation, setting the stage for a user-centric experience that goes beyond traditional methodologies.

The project extends its scope by dynamically generating content-rich slides that enhance both visual appeal and informational depth. The integration of Image Generation automates the process, ensuring that relevant visuals seamlessly align with the content of each slide. This harmonious blend of technology and creativity is a cornerstone of Slidester's scope, offering users a transformative experience where visual and textual elements coalesce effortlessly.

Crucially, Slidester liberates itself from temporal constraints, accessing the latest information from the internet to ensure presentations are not only visually engaging but also enriched with up-to-date and relevant data. This forward-looking approach reflects the project's commitment to adapting to the ever-evolving nature of information.

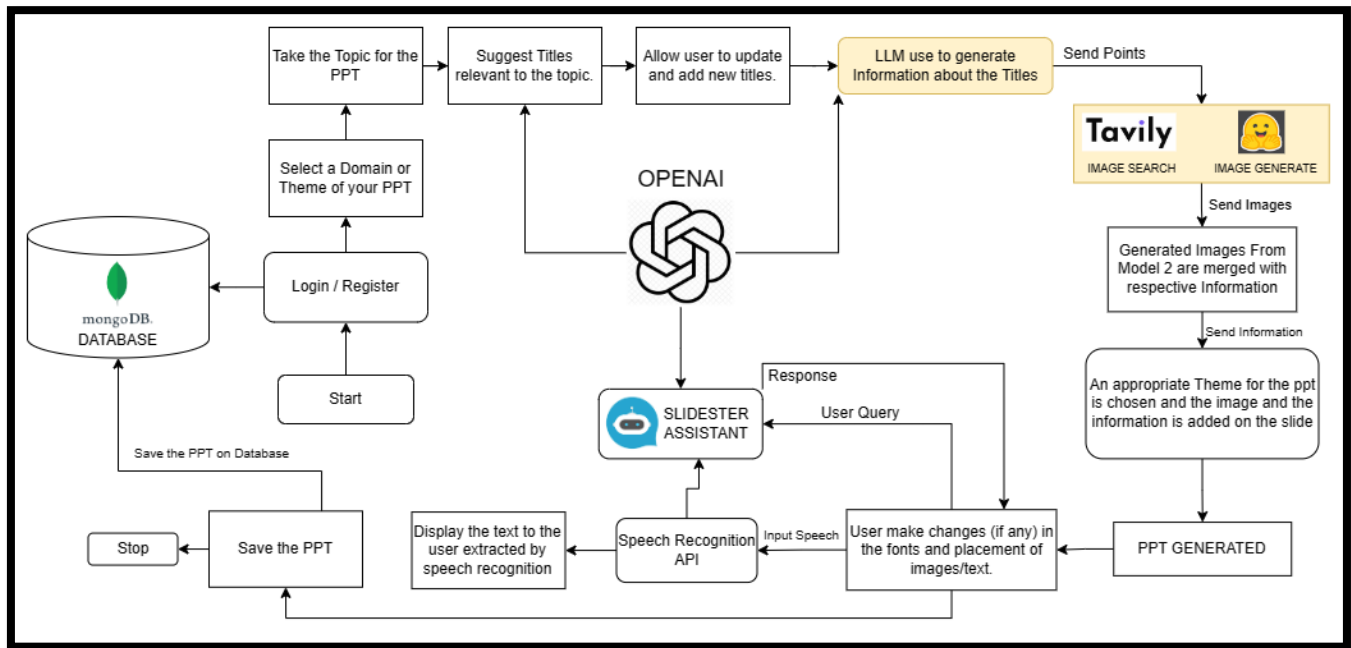
The scope of Slidester encompasses a commitment to real-time relevance through Internet-Powered Content, providing users with access to live data. This ensures that presentations remain current, reflecting the latest trends, statistics, and developments. The project's embrace of Document Integration further reinforces its adaptability to diverse information landscapes, acknowledging the importance of incorporating external data sources.

The project's versatile scope caters to a wide user base, including students, professionals, educators, and entrepreneurs. The provision of downloadable formats and creative templates adds a layer

of versatility, making Slidester an indispensable tool for individuals across diverse domains. This multifaceted approach positions Slidester as a groundbreaking solution emphasizing efficiency, creativity, and user-centric design in the realm of presentation creation.

The scope of Slidester extends beyond mere automation; it seeks to redefine the very fabric of presentation creation, promising an intuitive, enjoyable, and efficient experience for users across various domains. Slidester's ambitious scope sets it on a trajectory to be a transformative force in the technology-driven era of dynamic presentations

IV. Proposed Design (Architecture Diagram and other diagram)



4.1 ARCHITECTURE DIAGRAM

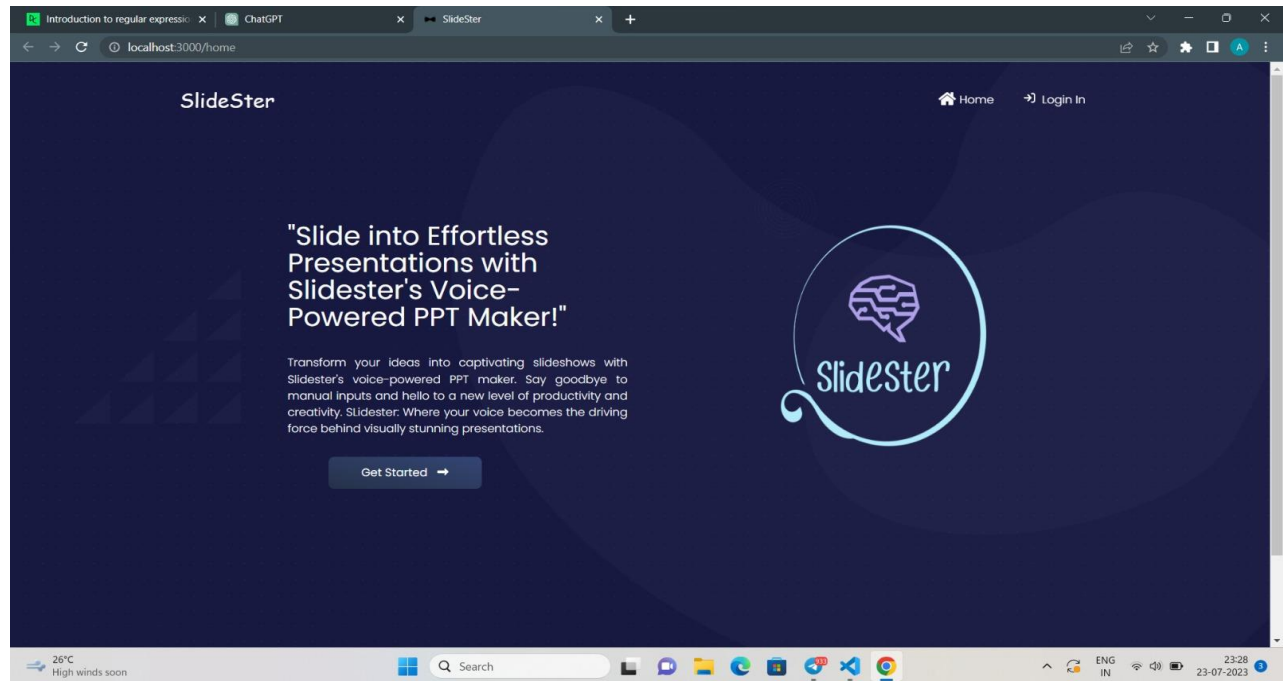
Explanation:

- Users initiate the process by either logging into their existing accounts or, for newcomers, registering on our website to create a personalized user profile.
- Once logged in, users navigate to the presentation configuration phase. Here, they select a thematic domain and specify a topic that aligns with the desired focus of their presentation.
- Our Language Model (LLM) takes center stage, suggesting preliminary slide titles based on the chosen domain and topic. Users are empowered to refine the structure further by specifying the number of points associated with each title.
- Users actively engage in the creative process by reviewing and, if necessary, editing the suggested titles. They can introduce entirely new titles or modify existing ones until satisfied with the composition.
- The finalized titles undergo a dual-model process for enrichment. Model 1 (LLM) generates detailed point information for each title, while Model 2 (Stable Diffusion Model) employs the

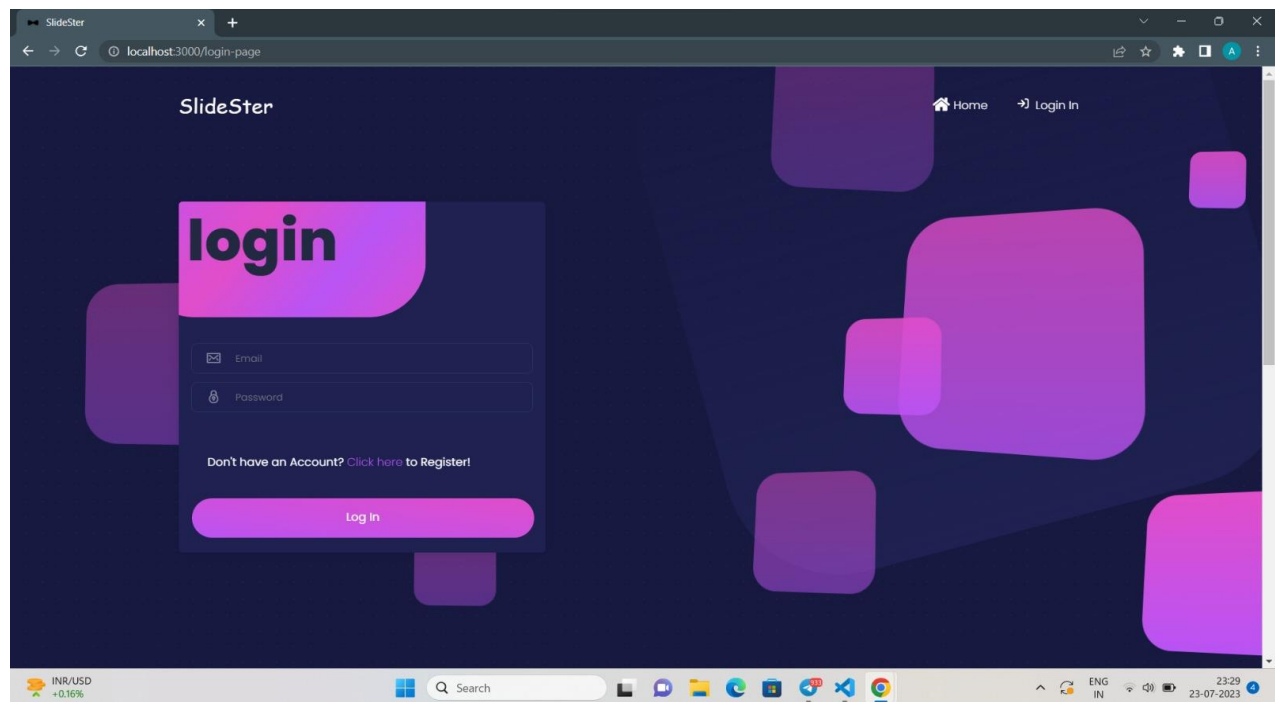
Latent Consistency Model (LCM-LoRA) to produce relevant images. Additionally, the presentation benefits from supplementary images sourced from the web through the Tavily client.

- The textual information and images collected in the previous steps seamlessly merge, resulting in a comprehensive presentation that combines well-crafted content with visually appealing elements.
- Users exercise control over the finer aspects of their presentation, including font styles and content modifications. This is facilitated through an intuitive conversational interface, allowing for a personalized touch.
- An interactive assistant, designed as a chatbot, becomes a valuable ally in the modification process. Users can interact with the assistant, providing necessary information for requested changes, which are then implemented seamlessly into the slides.
- The user has the convenience of exporting the final presentation in a preferred format, be it pptx or pdf, ensuring compatibility with different platforms and applications.
- All finalized presentations are automatically archived within the user's account. This feature serves as a practical repository, enabling users to conveniently revisit and reference their past presentations for future use or review.

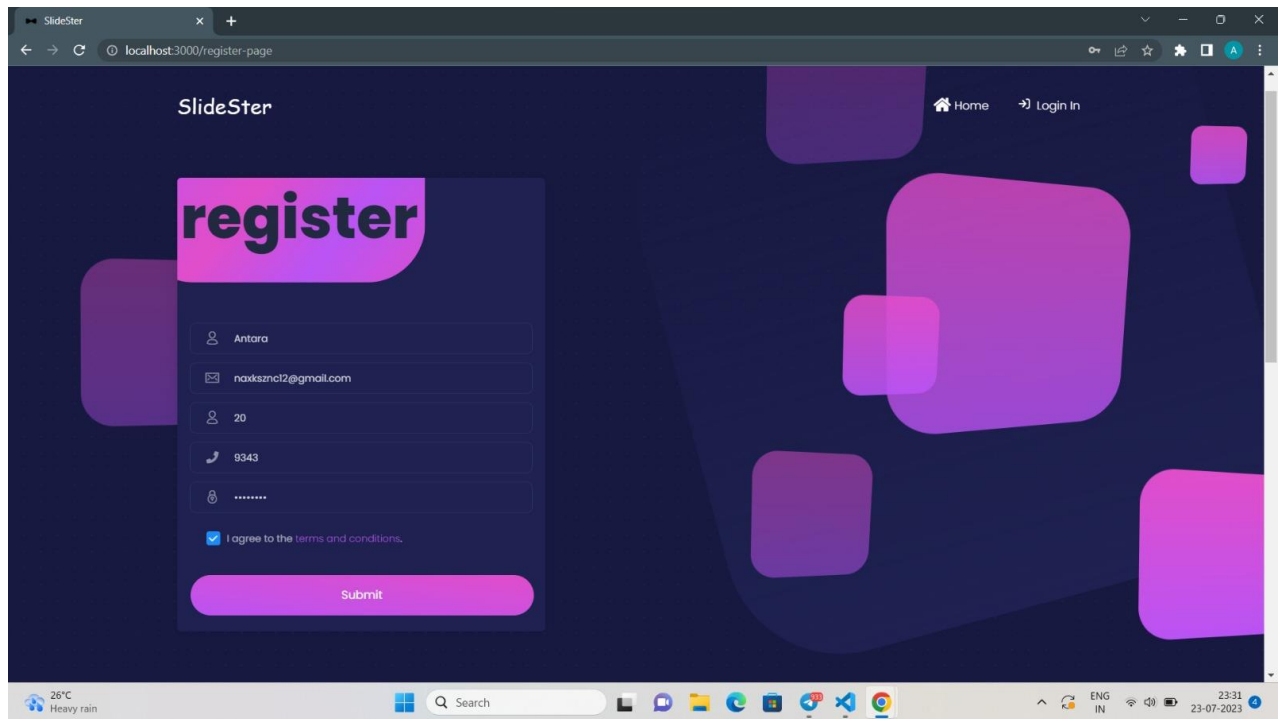
V. Results and Discussions



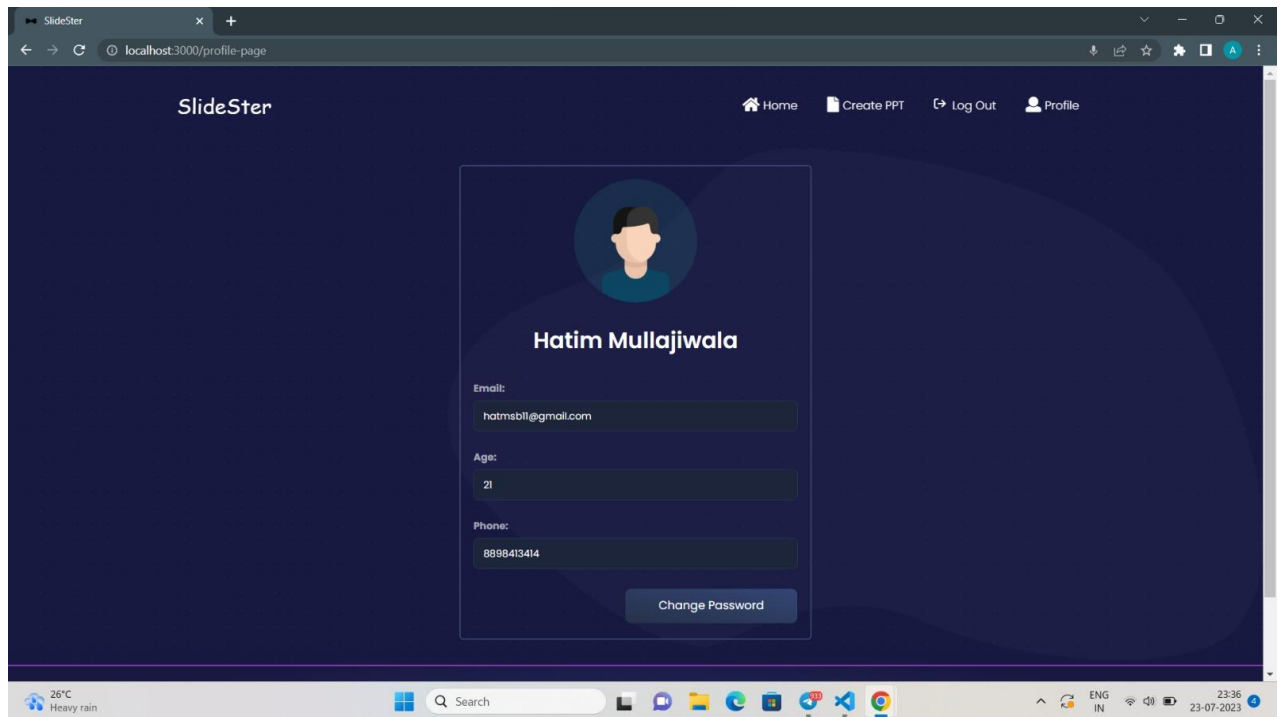
5.1) HOME PAGE



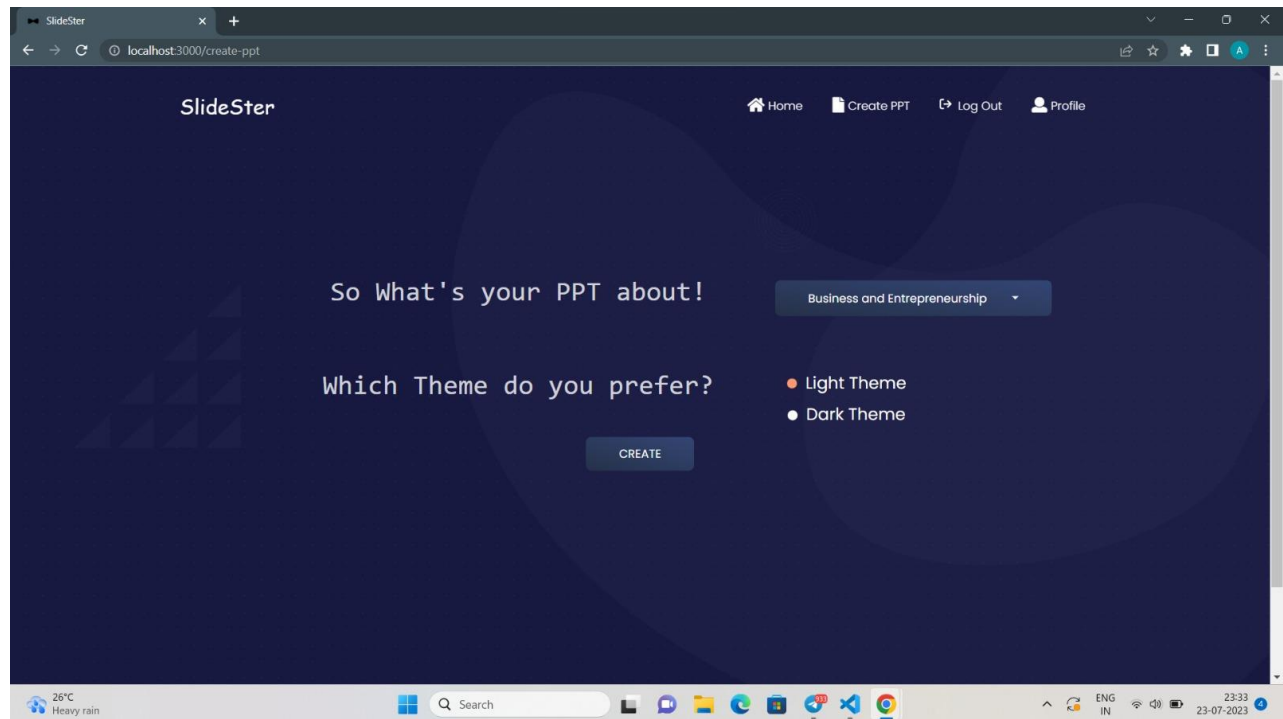
5.2) LOGIN PAGE



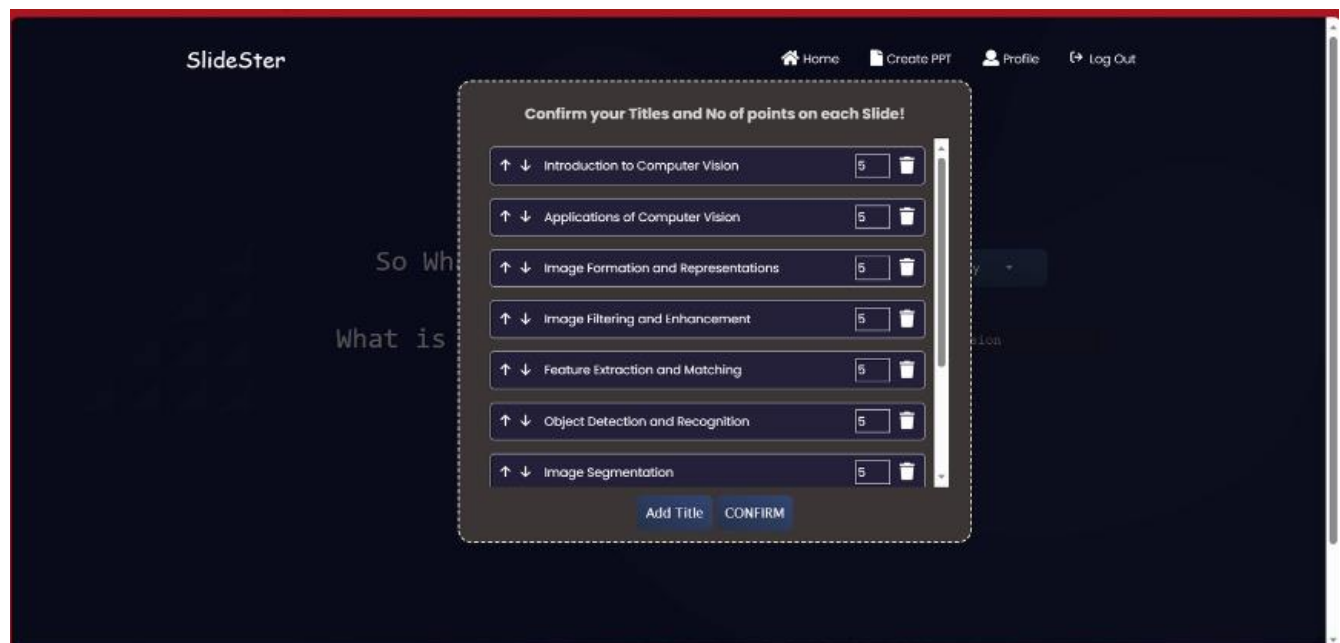
5.3) REGISTER PAGE



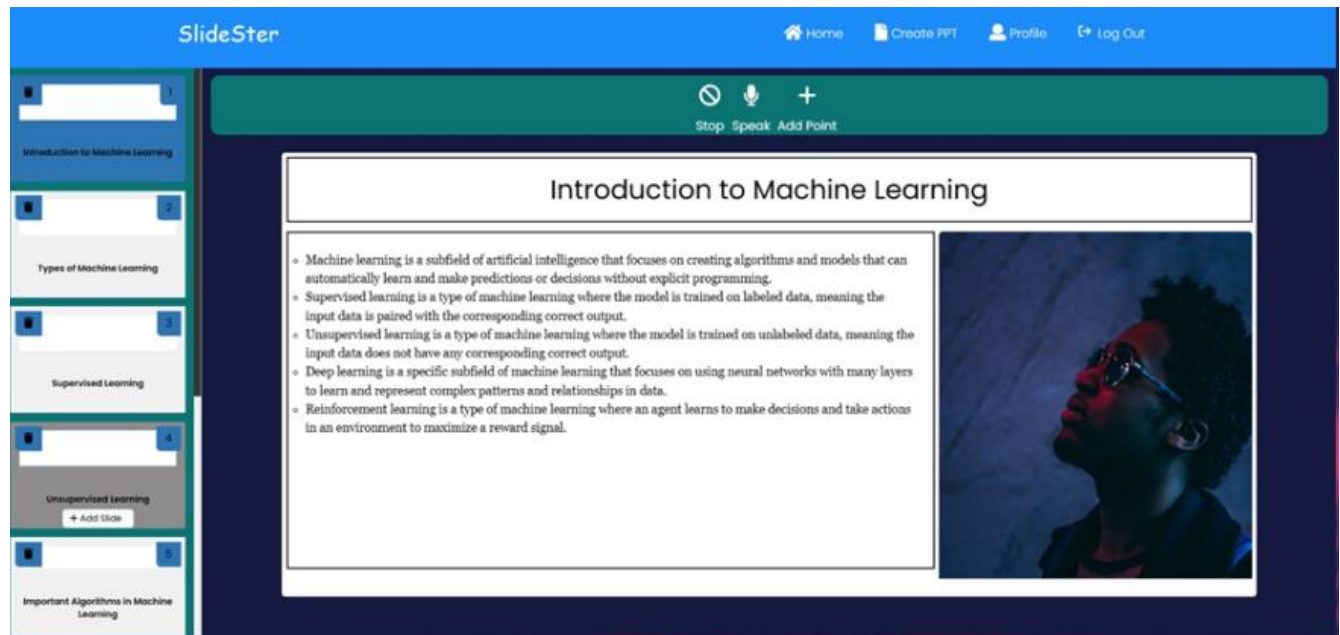
5.4) PROFILE PAGE



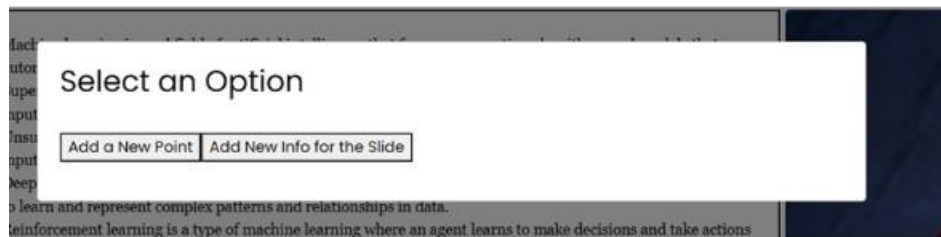
5.5) SUBJECT OF THE PPT



5.6) SELECT AND CONFIRM SUB-TITLES AND NO OF POINTS



5.7) PPT GENERATED



5.8) EDIT THE POINTS OR TITLES

VI. Conclusion

In conclusion, the Slidester represents a transformative and groundbreaking approach to presentation creation. By harnessing the power of advanced technologies, including NLP algorithms and speech recognition, the project offers users a seamless and intuitive means of converting spoken content into visually stunning presentations. With automated image generation and a user-friendly editing toolbar, Slidester empowers individuals from diverse domains to unleash their creativity and deliver captivating messages without the burden of tedious manual tasks. The project's scope extends to a wide range of users, from students to professionals, providing a versatile and dynamic platform for impactful communication. With its versatile and dynamic platform, the project opens doors for more efficient and enjoyable presentation-making experiences. Slidester paves the way for hassle-free presentation-making, saving time, and elevating the quality of presentations, leaving a lasting impression on audiences worldwide.

VIII. References

- [1] Sharma, N., & Sardana, S. (2016). A real time speech to text conversion system using bidirectional Kalman filter in Matlab. *Advances in Computing and Communications*. <https://doi.org/10.1109/icacci.2016.7732406>
- [2] Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C. J., Casper, J., Catanzaro, B., Cheng, Q., Chen, G., Chen, J., Chen, J., Chen, Z., Chrzanowski, M., Coates, A., Diamos, G., Ding, K., Du, N., Elsen, E., . . . Zhu, Z. (2016). Deep speech 2: end-to-end speech recognition in English and mandarin. *International Conference on Machine Learning*, 173–182. <https://arxiv.org/pdf/1512.02595>
- [3] Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S. K. S., Ayan, B. K., Mahdavi, S. R., Lopes, R. G., Salimans, T., Ho, J. C., Fleet, D. J., & Norouzi, M. (2022). Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. *ArXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2205.11487>
- [4] Belote, P., Bidwai, S., Jadhav, S. R., Kapadnis, P., & Sharma, N. (2017b). Enhancing Automatic PPT Generation Technique through NLP for Textual Data. *Nternational Journal of Advanced Research in Computer and Communication Engineering*. <https://doi.org/10.17148/ijarcc.2017.63237>
- [5] Kao, O., & Joubert, G. R. (2000). A content-based Internet search engine for analysis and archival of MPEG-1 compressed newsfeeds. *International Conference on Multimedia and Expo*. <https://doi.org/10.1109/icme.2000.871565>
- [6] P. M., Santra, S., Bhowmick, S., Paul, A., Chatterjee, P., & Deyasi, A. (2018). Development of GUI for text-to-speech recognition using natural language processing. 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech). <https://doi.org/10.1109/iementech.2018.8465238>
- [7] Bhat, A., Rao, A. C., Bhaskar, A., Adithya, V., & Pratiba, D. (2018). A cost-effective audio-visual summarizer for summarization of presentations and seminars. 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS). <https://doi.org/10.1109/csitss.2018.8768740>
- [8] Ramzan, S., Iqbal, M. M., & Kalsum, T. (2022). Text-to-image generation using Deep Learning. *IEEC 2022*. <https://doi.org/10.3390/engproc2022020016>