

Industrial training presentation



Name : Somdutt Choudhary

Roll no:22EAIAD200

Branch: AI & DS

Collage: Arya collage of engineering

Table of Contents

- Company Details
- My Roles and Responsibilities as Intern
- TextToVideo
- Technology Stack
- Workflow
- Code Snippets
- Challenges and Solutions
- Key Learnings
- Project Conclusion
- Thank You!



IceHut

Delivering Digital Transformation

Ice Hut Technologies PVT LTD

Analysis of data

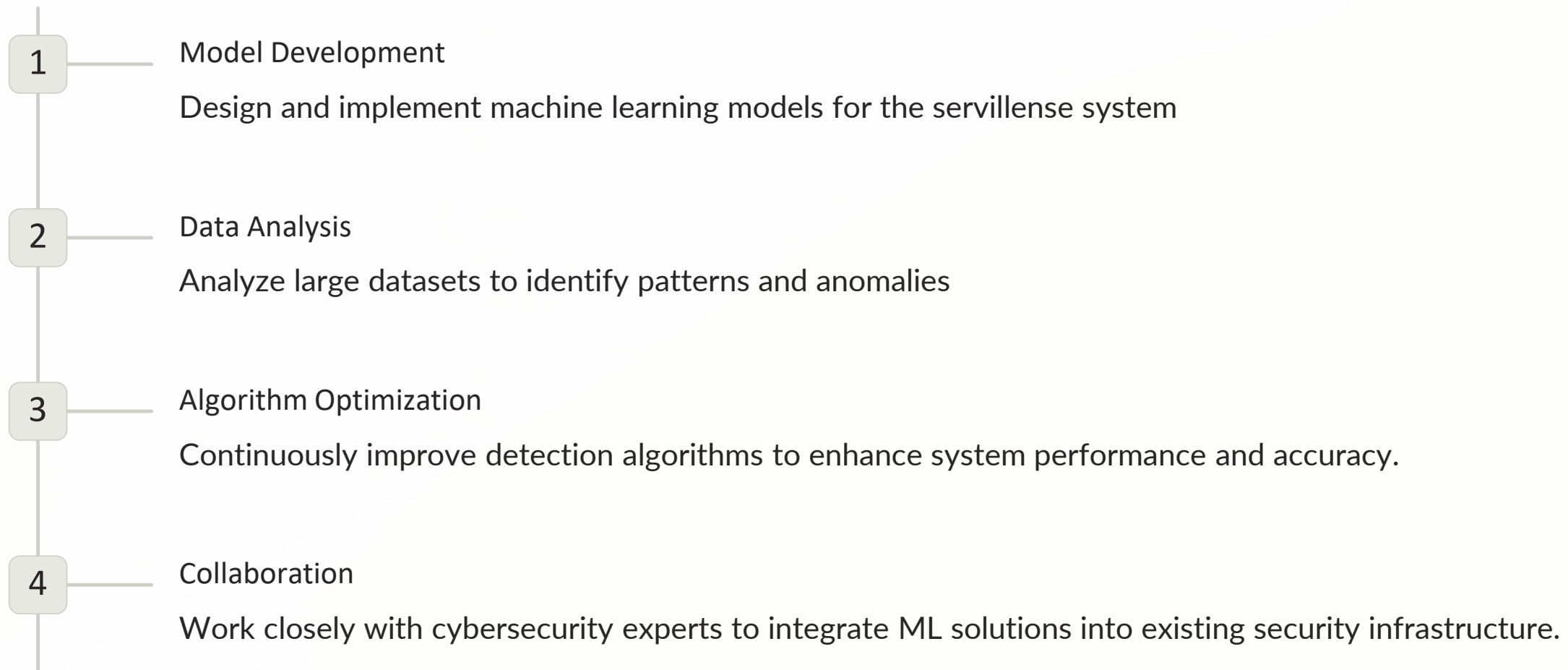
Location

Office 4, Radhika Vihar D, Near Iskcon Temple, Mansarovar,
Jaipur 302020

Focus

creating ai based security, and surveillance systems

My Roles and Responsibilities as intern

- 
- 1** — Model Development
Design and implement machine learning models for the servillense system
 - 2** — Data Analysis
Analyze large datasets to identify patterns and anomalies
 - 3** — Algorithm Optimization
Continuously improve detection algorithms to enhance system performance and accuracy.
 - 4** — Collaboration
Work closely with cybersecurity experts to integrate ML solutions into existing security infrastructure.

TextToVideo:

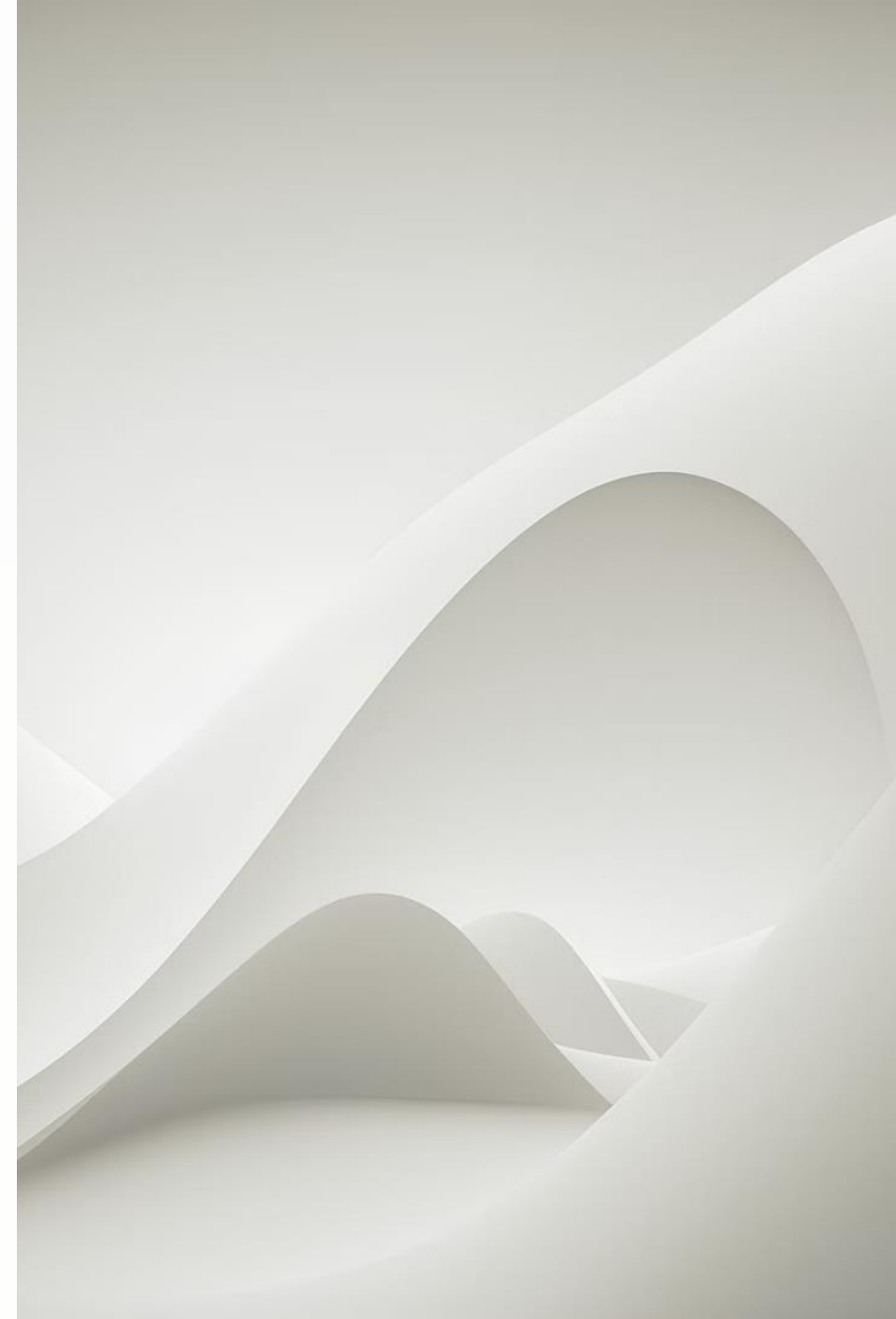
project overview

TextToVideo is an AI-powered tool that simplifies video creation by leveraging GPT-4 to generate engaging scripts. You can customize these scripts using various text-to-speech engines, including pyttsx3, ElevenLabs, gTTS, and OpenAI TTS.

Create unique videos with TextToVideo by utilizing its stock footage library or uploading your own custom clips.

Key features of TextToVideo include:

- **GPT-4 Powered Script Generation:** Creates engaging scripts based on user input.
- **Text-to-Speech Engines:** Brings the scripts to life with customizable voices.
- **Visual Library:** Offers a wide range of stock footage for seamless video creation.
- **Custom Clip Upload:** Allows for personalized visual elements.



Technology Stack



MoviePy

A Python library for video editing and manipulation.



Pillow

A Python Imaging Library (PIL) fork used for image processing and manipulation.



OpenAI

A powerful AI platform used for text generation, image generation, and more.



ElevenLabs

A text-to-speech platform for generating high-quality audio.



Numpy

A fundamental Python library for numerical computing and data analysis.



gTTS

A Python library for converting text to speech using Google's Text-to-Speech API.



pyttsx3

A Python library for text-to-speech conversion.



OpenAI Whisper

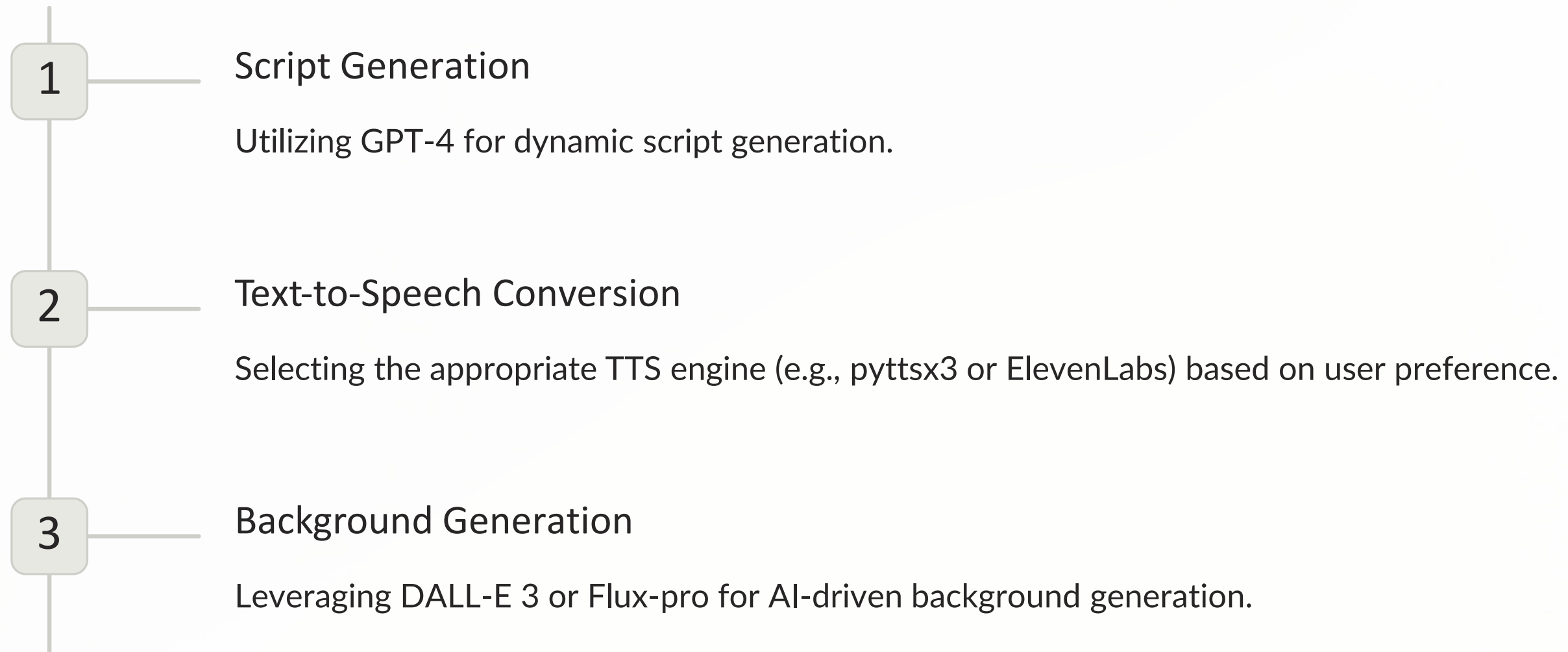
A powerful automatic speech recognition system developed by OpenAI.



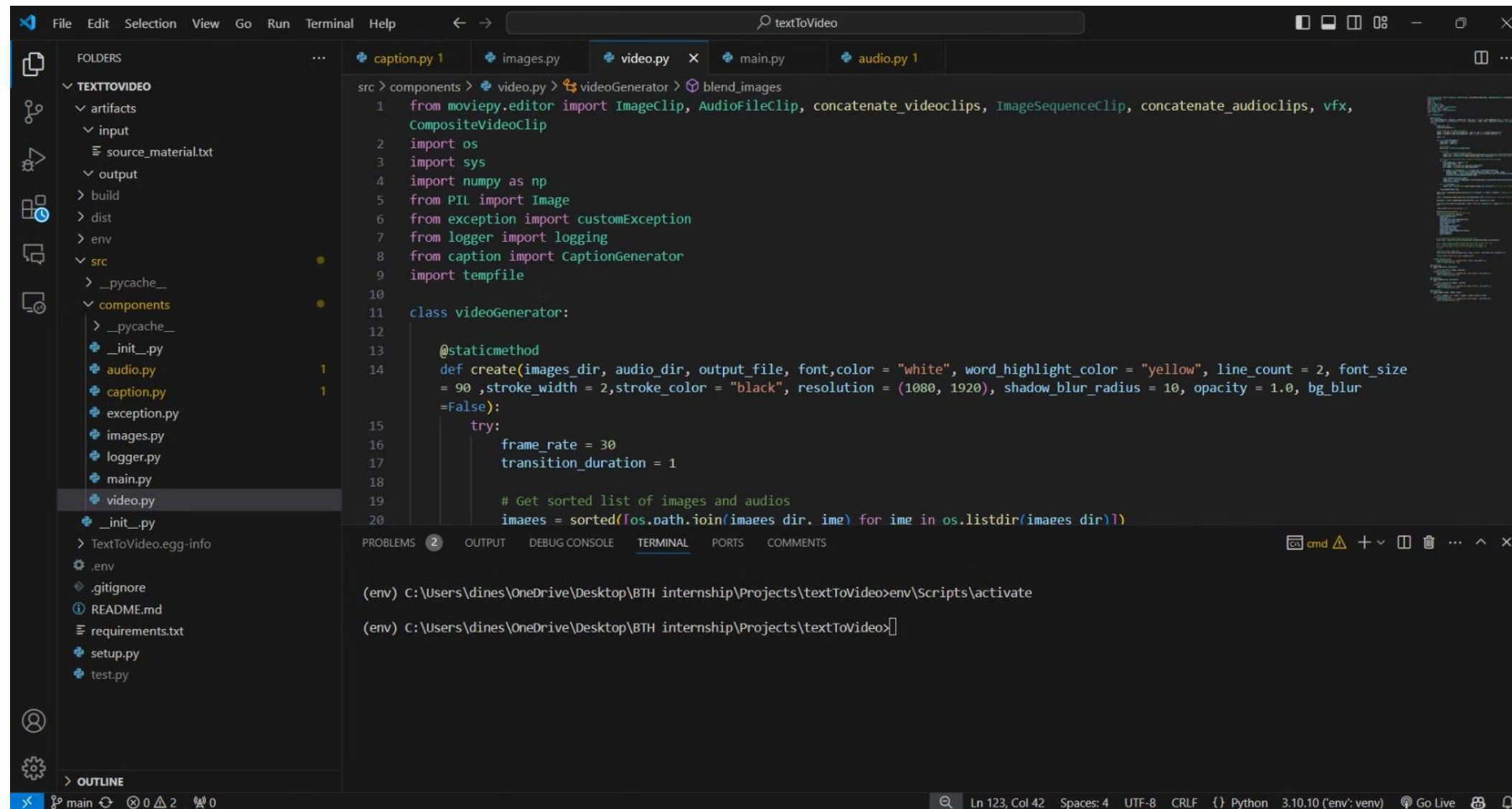
Replicate

A platform that allows users to run and deploy machine learning models.

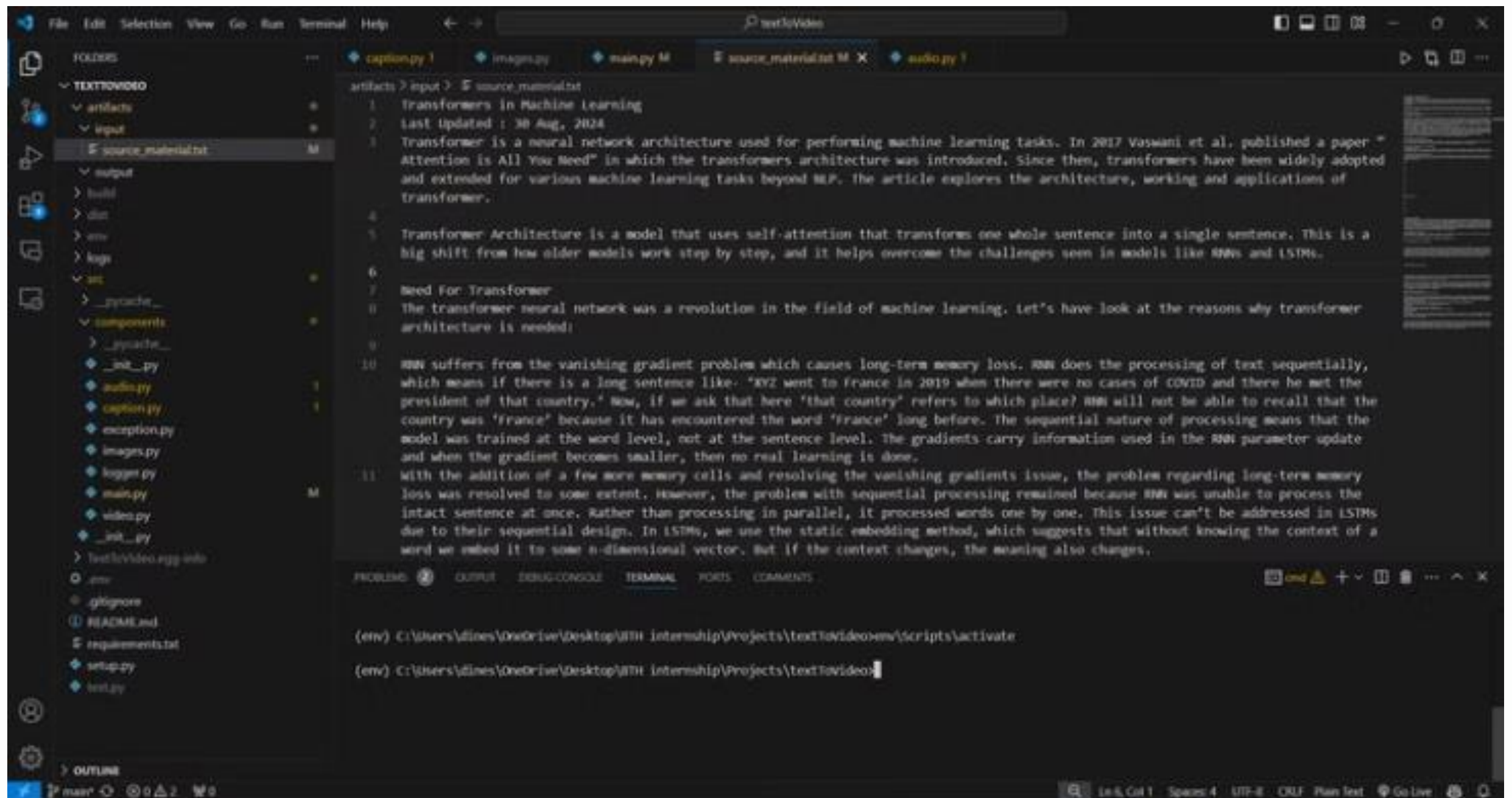
Workflow



Code Snippets:

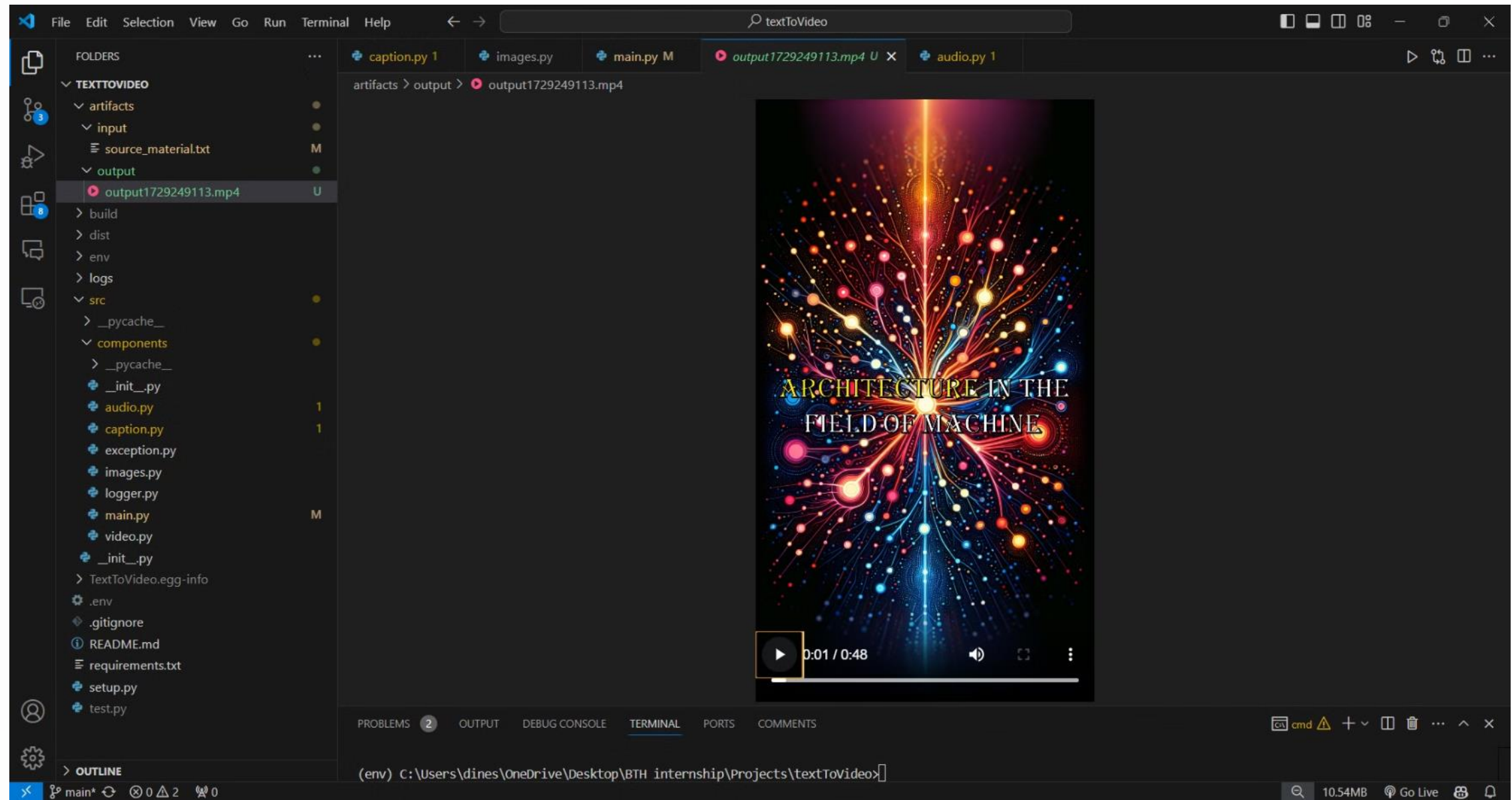


glimps of file system and code snippets



OUTPUT :

generated video output



An abstract graphic on the left side of the slide, consisting of multiple layers of white, wavy, paper-like shapes that create a sense of depth and movement. The layers are stacked and overlap, with some edges pointing towards the viewer and others receding.

Challenges and Solutions

Challenge

Creating dynamic captions that sync with the video and voiceover.

Solution

Implementing OpenAI Whisper for accurate, real-time transcription and caption generation.

Key Learnings

1 Video Editing Expertise

Gained in-depth knowledge of video editing libraries like MoviePy.

2 AI Tool Mastery

Mastered working with AI-based tools such as GPT-4, OpenAI Whisper, and ElevenLabs.

3 Automated Captioning

Learned how to automate the process of captioning and synchronizing it with audio.

4 Database Operations

Acquired hands-on experience with MongoDB and programmatic CRUD operations.





Project Conclusion

Conclusion

Summary: Successfully developed a tool that automates video creation from text inputs using advanced AI techniques.

Future Enhancements

Focus on improving the accuracy of captioning and expanding the tool to support more customization features for users.

Thank You!

We appreciate your time and feedback. We hope you enjoyed this presentation.