What would have happened if 15 year old Taylor Swift had AI ?

- Singer / Music creator
- Young
- Talented
- Limited resources
- Proof of concept

2024

Deep Learning Dudes Studio

DLD-Vision

-Effortlessly Create Dynamic MV

Group members: Ruixiang Wang
Fernando Ramirez Gonzalez
Hengsheng Li
Minjun Zhong

Contents

01

Record of the Project Process

Task breakdown

Development Phase

01 Divide the generation into 3 parts: lyrics, video and music.

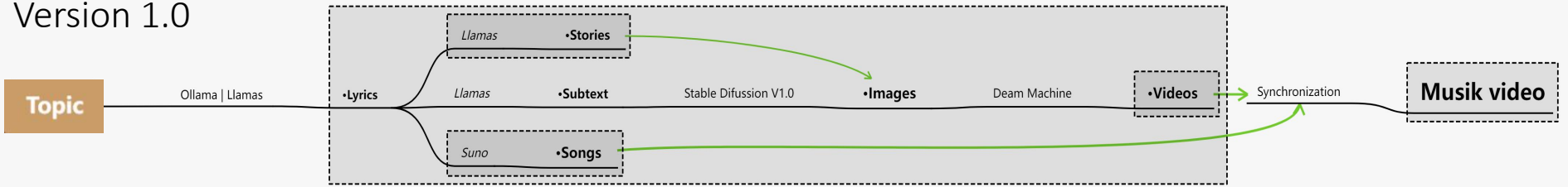
02 Define a workflow and assign the tasks

03 Integrate all modules, fix bugs and do documentation

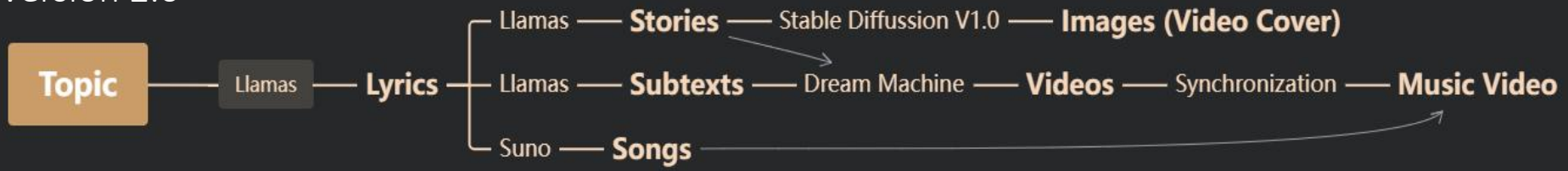
04 Enhancements and improvements made to the project content (Bonus)

Lets automate, our workflow

Version 1.0



Version 2.0



Task assignment and group work

Ruixiang Wang

* Design Workflow
* Presentation summary
* Identify models APIs
* Test Prompts and Videos
* Edit MV manually
* Coordinate team and keep track progress

Fernando

* Design Workflow
* Work And Test on TTS | BARK \|LLM|SUNO
* Testing Singing Voice
* Test Prompts and Videos
* Regex for LLM text extraction
* Video generation

Hengsheng Li

* Implement Suno API
* Test song generation
* Test final video generation
* Test all code and fix bugs
* Test code deployment.

Minjun Zhong

* Text to videos
* Work And Test on Llama 3|DreamMachine
* Identify models APIs
* Document record

Lyric generation

- Used LLama with the template and some regex

Key takeouts for the prompt that are useful later:

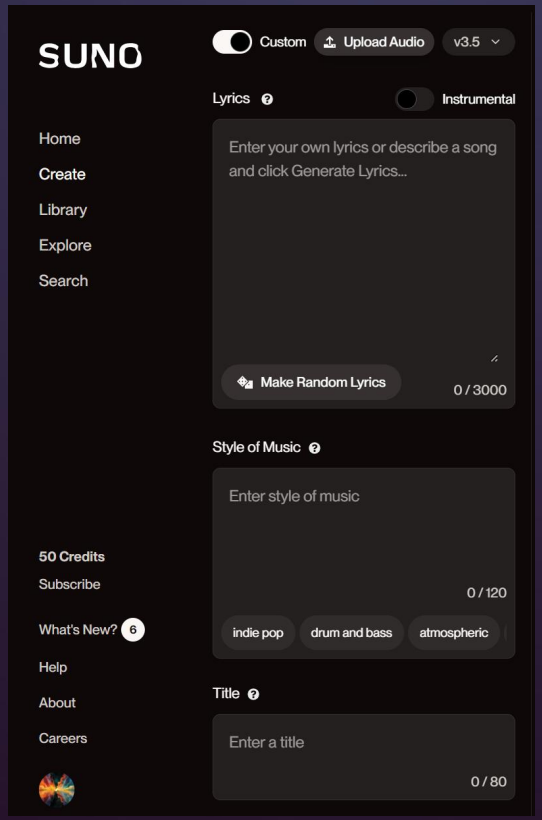
* BPM
* Tempo
* Key
* Structured lyric format, feed directly
* Vocal performance as tags later
* Physical actions used as subtexts later

DeepLearningDudes_Music_Video [WSL: Ubuntu] - LLM_processing.py

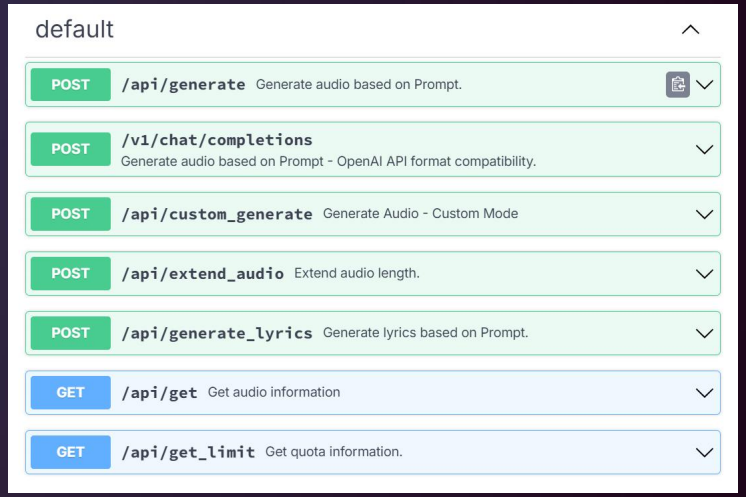
```
Example:
                **Song Title:** "Swipe Right on Memories"

                **Genre:** Pop-Rock with a hint of Electronic elements

                **Tempo:** Moderate (around 120 BPM)

                **Time Signature:** 4/4

                **Key:** C Major

                **Lyrics:**

                [Verse 1]
                I was scrolling through my feed, feeling so alone
                When I saw your face, and my heart started to moan
                A swipe right, and our stories aligned
                Little did I know, our love would be redefined

                [Verse 2]
```

Suno Integration

WebUI



Request



Local Host

DeepLearningDudes_Music_Video [WSL: Ubuntu] - suno_api.py

```python
def custom_generate_audio(payload):
    url = f"{base_url}/api/custom_generate"
    response = requests.post(
        url, json=payload, headers={"Content-Type": "application/json"}
    )
    print(f"Response text: {response.text}")
    return response.json()
```

Video Prompt Generation

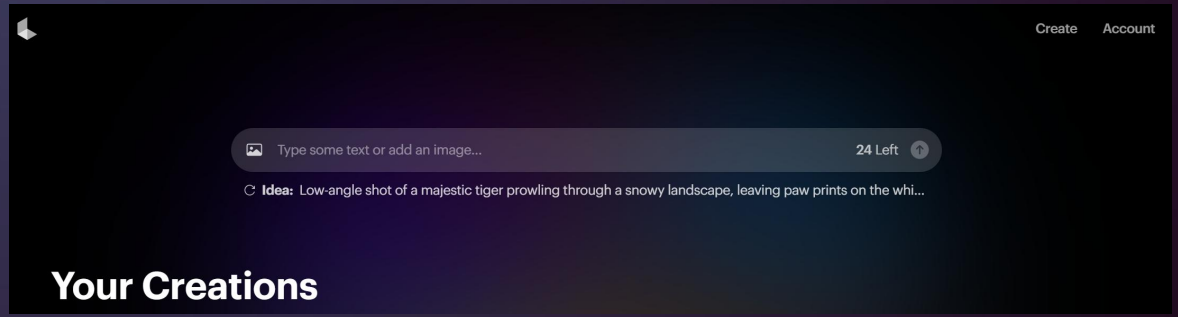
- Physical scenario
- Character description
- Character actions
- Plot development
- Negative prompt

DeepLearningDudes_Music_Video [WSL: Ubuntu] - LLM_processing.py

```
Template:
        *Scene 1: Introduction*
        Prompt:
        Physical scenario: [description]
        Character/s physical description: [name][man/woman][age][hair color][hair length][hair style][eye color][face descriptions][height][clothing style][clothing color][clothing material], [name][man/woman][ag
        Character actions: [description]
        Plot development: [description]
        Negative prompt: bad anatomy, bad proportions, blurry, cloned face, deformed, disfigured, duplicate, extra arms, extra fingers, extra limbs, extra legs, fused fingers, gross proportions, long neck, malfor
        *Scene 2: scene 2 title*
        Prompt:
        Physical scenario: [description]
        Character/s physical description: [name][man/woman][age][hair color][hair length][hair style][eye color][face descriptions][height][clothing style][clothing color][clothing material], [name][man/woman][ag
        Characters actions: [description]
        Plot development: [description]
        Negative prompt: bad anatomy, bad proportions, blurry, cloned face, deformed, disfigured, duplicate, extra arms, extra fingers, extra limbs, extra legs, fused fingers, gross proportions, long neck, malfor
```

Luma Integration

WebUI



Function

DeepLearningDudes_Music_Video [WSL: Ubuntu] - util.py

```
1 def dreamMachineMake(prompt, access_token, img_file):
2     url = "https://internal-api.virginia.labs.lumalabs.ai/api/photon/v1/generations/"
3
```

Integration

- Join all videos and music
- Video length is fixed
- Music is not time fixed for all parts of the text
- Synchronization is poor when automated
- Editing and synchronization should be done manually

Overview of models used during development

| Title to lyrics
| Lyrics to stories |

* **Chatgpt**
* **Prompt perfect**
* **Llama3**

| Lyrics to songs |

* **Suno**

| Stories to images |
| Stories to videos |

* **Llama3**
* **DreamMachine (Luma)**
* **Stable Difussion (v1.0)**

| Synchronisation |

* **Completed Manually**

Discarded models and the reasons

| Title to lyrics
| Lyrics to stories |

None

| Lyrics to songs |

*** Gen Music**

*** GTT(Bark and Googles model)**

*** TTS**

| Stories to images |
| Stories to videos |

Realistic

| Synchronisation |

*** Automatical synchronisation**

Enhancements and improvements made to the project content (Bonus)

01 Outputs are prompt controlled

02 Creating dialogue enriched with subtexts

03 Image style is template controlled

04 Characters in multiple video clips are largely alike

P A R T T W O

02

Presentation of Project Results


Swipe right on memory
BY | DLD Studio

Advantages and improvements achieved through the project

01

- End-to-end automated pipeline (use of JSON)

02

- Use positive and also negative prompts during image generation to ensure that the characters' limbs and expressions appear natural

03

- When generating the video, we used the subtext of the lyrics to ensure the video's stability and its relevance to the lyrics.

Advantages and improvements achieved through the project

04

- Control over individual keyframes (easilly replacebale)

05

- Use of examples for structuring the LLM output.

06

- We enhance coding efficiency by simplifying and updating the workflow.

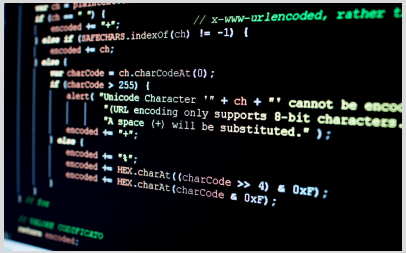
03

End of Project Reflections

Insights Gained from Project Learning

01

Enhanced programming abilities and code problem-solving skills.



02

Deepened understanding of the process of building artificial intelligence models.

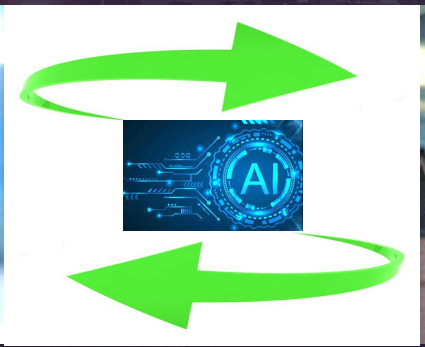
03

learned how to effectively collaborate and communicate with team members.

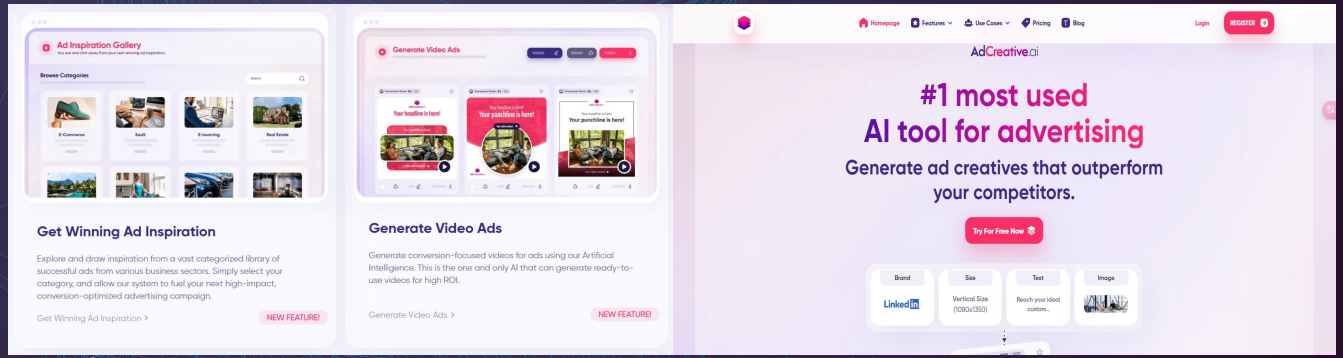
Other possible applications

GET STARTED

- How to use AI to generate high-quality advertising videos ?



Adcreative.ai :



Can only generate Videos

Our "Adsvision"

-Effortlessly Create Dynamic Ad Music Videos

Thanks for watching !

2024

Deep Learning Dudes

Group members: Ruixiang Wang | Fernando Ramirez Gonzalez | Hengsheng Li | Minjun Zhong